UNIVERSITY OF TARTU
School of Economics and Business Administration

Aruzhan Taufikova
Gulustan Ismayilova

GENDER DIVERSITY OF WORKFORCE AT FIRMS AND INNOVATION PERFORMANCE: EVIDENCE FROM FIRM-LEVEL DATA IN SPAIN

Master's Thesis

Supervisor: Priit Vahter, Professor of Applied Economics

Tartu 2023

We have written this Master's Thesis independently. Any ideas or data taken from other authors or other sources have been fully referenced.

Abstract

Gender diversity is often found to have a beneficial impact on the firm performance. Nevertheless, the studies on the link between gender diversity and innovation inputs and outputs are scarce. This paper examines whether increased gender diversity at firms fosters product and process innovation. Furthermore, as a key novelty of the paper, the relationship between gender diversity and open innovation, which refers to collaboration with external sources, is investigated. The empirical analysis is conducted on the firm-level panel PITEC data from Spain during 2004-2016 using propensity score matching technique. We find that firms with less than 10% increase in share of women in the total workforce are more likely to engage in product, process, and open innovations. However, over a 3-year period a negative association has been observed after a 10% increase in share of women and open innovation, as well as a negative relationship between decreased gender diversity and open innovation. Our results support the presence of a nonlinear relationship.

Keywords: gender diversity, innovation performance, open innovation, Spain

CERCS code: S180

Introduction

The growing participation of women in the labour force revealed the importance of gender diversity in the workplace and opened a new field for researchers to study. The economic empowerment of women is linked to increased access to education, changing societal attitudes towards gender roles, economic drivers, and government initiatives aimed at reducing gender disparities. While Norway was the first country to legislate gender quotas for corporate boards in 2003 (Comi et al., 2020), Spain, known for having a historically low female participation in the workforce, was the second country to implement quotas for 40% female representation on boards in 2007 (Reguera-Alvarado et al., 2017; Yang et al., 2019).

The aim of this paper is to examine the relationship between gender diversity and the innovation performance of firms in the context of Spain. Starting with the association between increased share of women in the total workforce and product and process innovation, the paper will expand further to explore the link between gender diversity and open innovation indicators.

Innovation refers to the introduction of a new or improved product or process and implementation of a new marketing or organizational methods (OECD & Eurostat, 2005). It is used as a tool for boosting economic growth, attaining a competitive advantage, and enhancing the financial performance (Alsos et al., 2013; Tahir et al., 2021). As innovation requires communication, collaboration, and creativity to tackle complex tasks, it is frequently linked to gender diversity (Ostergaard et al., 2011). In this study, we define gender diversity at workforce as a balanced representation of men and women across different positions and departments of a company.

Nowadays, innovative companies have adopted an "open innovation" framework, where organizations employ both internal and external sources (Chesbrough, 2003). Prior studies suggest the existence of complementarity between internal R&D and external knowledge sourcing (Cassiman and Veugelers, 2006; Love et al., 2014; Schmiedeberg, 2008). Although the studies on innovation and gender are limited, there is even less research available examining the relationship between gender diversity and indicators of open innovation. Given the importance of open innovation to many industries, we address the relationship between gender diversity and open innovation to fill this research gap.

Empirical analysis is based on a PITEC panel firm-level dataset. We will use the propensity score matching technique to examine the difference between companies with less and more than 10% threshold increase in the proportion of women and those without these characteristics.

Our findings show that companies with an increase in the share of women up to 10% have a positive association with product, process, and open innovation. However, increase above the threshold over a three-year period decreases utilization of open innovation by 1.7%. Another interesting outcome is the negative association between reduced gender diversity indicator and open innovation by 1.3%. Results show some support for the presence of a nonlinearity, suggesting an inverted U-shaped relationship. This study has several contributions to the literature. We investigate the role of women representation by analyzing both an increase and decrease in share of women, as well as gender diversity index. Secondly, we conduct the analysis on the total workforce and find some evidence of a nonlinear relationship, which has important policy implications.

The rest of this research is structured in the following way. Section 1 presents a literature review regarding the existing studies about gender and innovation. This section is divided into theoretical considerations and empirical results. Section 2 is focused on empirical analysis. We describe the data, variables, and the propensity score matching methodology. This is followed by the discussion of results, key findings, sensitivity analysis, and policy implications in section 3. Finally, the conclusion summarizes the study, discusses limitations, and considers ideas for further research.

1. Literature review

1.1 Theoretical evidence

Firms undertake innovative activities to improve their overall performance. Existing studies on gender diversity are focused mainly on the impact on firm performance (Díaz-García et al., 2013). Innovation is defined as the successful application of a new or improved product or process, implementation of a new marketing or organizational method in the influential Oslo Manual, which provides the framework for research on innovation and is used in innovation surveys (OECD & Statistical Office of the European Communities, 2005). In recent literature, there are two types of innovation distinguished: product and business process innovations (OECD & Eurostat, 2018).

Existing literature on gender diversity in the workplace mainly follows two different views. The first “value-in-diversity" approach states that gender diversity positively affects innovation performance by providing different creative ideas, broader perspective and competence base, which yields social and informational benefits (Ostergaard et al., 2011; Xie et al., 2020). On the other hand, social identity theory points out that in the existence of gender diversity, formation of groups with similar interests causing social exclusion can lead to conflicts and decrease in team performance (Díaz-García et al., 2013; Tsui et al., 1992). However, since innovation takes place in knowledge-based working environments where individuals execute challenging tasks, and different points of views, fresh perspectives are highly valued, conflicts may not be an obstacle (Xie et al., 2020). Dezso and Ross (2012) show that while homogeneous teams work better on simpler tasks, heterogeneous teams perform better on more complex ones. Therefore, it can be suggested that value-in-diversity theory overshadows social identity theory in case of complex tasks.

1.1.1. Open innovation

The concept of open innovation has gained significant attention both in academia and industry recently. The initial premise of the open innovation stems from the notion that a firm needs to collaborate with external factors to innovate and remain competitive (Chesbrough, 2003; Dahlander & Gann, 2010; Laursen & Salter, 2006). According to Durst and Ståhle (2013), the high cost of innovation and the need to develop products quickly requires collaboration with internal and external actors, even with competitors. Over the last two decades, the definition of open innovation term has evolved multiple times, leading to ongoing conceptual ambiguity (West et al., 2014). Chesbrough (2003, p.24) defines open innovation as “a paradigm that assumes that firms can and should use external as well as internal ideas, internal and external paths to market, as firms look to advance their technology”. Later, the intentionality of outflows and inflows of knowledge was also recognized under the open innovation term (Chesbrough et al., 2006). In our paper, we define the open innovation term as the combination of internal R&D and external knowledge sourcing, following the Love et al. (2014) approach.

Laursen and Salter (2006) relate openness to the quantity of external sources, whereas Henkel (2006, as cited in Dahlander & Gann, 2010) describes openness as a means of exposing organizations’ hidden ideas. Dahlander and Gann (2010) suggest two types of innovation:

inbound and outbound, where former refers to acquiring and sourcing knowledge and outbound form to revealing and selling the expertise.
Some studies argue that possible disadvantages of open innovation are not given enough attention or overlooked with greater focus being placed on the benefits it provides. For instance, Laursen and Salter (2006) show that while intellectual property rights can incentivize innovation, they can also impede the process of building upon existing innovation, creating barriers to entry. On the contrary, Dahlander and Gann (2010) argue that exposing resources to other companies and being an open company can make it difficult to protect intellectual property rights and obtain economic benefits.

The systematic analysis of 701 papers on open innovation revealed the importance of the role of individuals involved in the innovation processes, as one of the factors promoting open innovation (Durst & Ståhle, 2013). As an example, Lindegaard (2010, as cited in Durst & Ståhle, 2013) emphasizes the significance for individuals to possess the certain communication skills and mindset in driving open innovation processes. When exploring it through the gender factor, collaboration and competitiveness suggest different implementations for open innovation. Bertrand (2011) studied psychological differences between men and women, and noted that, on average, women tend to be more risk-averse and social minded, while men behave more competitively. In regard to negotiations, evidence shows that depending on a situation, women tend to negotiate better for other people rather than for themselves, while men negotiate similarly in both cases. Hence, it can be argued that women are more open to collaboration than to competition and thus may perform well in open innovation cases. Previous research findings in the fields of economics and psychology are also consistent with the notion that men tend to take more risks than women. (Croson & Gneezy, 2009).

1.2 Empirical evidence

1.2.1 Gender diversity and innovation

Even though most of studies have mainly focused on researching the different indicators of innovation rather than the role of its actors, recent papers examine the correlation between employee diversity and innovation (Alsos et al., 2013; Ritter-Hayashi et al., 2019). Those papers have centered on the innovation indicators as input, output and efficiency. The existing theories suggest that gender diversity has a positive impact on innovation. The study by Østergaard et al. (2011) determines that gender and education diversity in Danish companies positively and significantly affects the likelihood to innovate. Research by Xie et al. (2020) examines the relationship between the share of women in R&D team and the innovation efficiency of the firm, for which an industry level data from China was used. The evidence shows that the higher the task intensity and market competitiveness, the higher the positive effect of gender diversity on innovation efficiency (Xie et al., 2020).

Applying an institutional framework in a meta study of 1069 firms in 35 countries, Zhang (2020) observed that the association between gender diversity and firm performance can be influenced by both normative and regulatory standards so that if gender diversity is viewed as a conventional practice in a particular country or industry, the higher the likelihood to have a positive correlation. The following subsections offer a review of literature focused on the topic of women on corporate boards, gender diversity in the Spanish firms, and connection of gender diversity and open innovation. Subsequently, **Table 1** is presented in the appendix and provides a summary of the literature review, including a description of the data, the methodology employed, and the corresponding results.

1.2.2 Gender diversity in corporate boards

The empirical evidence regarding the connection between gender diversity on corporate boards and firm performance is inconclusive, showing positive, negative, or non-linear relationship (Gong & Girma, 2021). When examined in managerial level, according to Griffin et al. (2021), introducing compulsory quotas for women in the board management positions raises awareness, encourages female involvement in the labour market and affects positively on the firm value. Nonetheless, existing empirical studies show mixed results regarding the impact of these quotas on firm performance (Yang et al., 2019).

On the one hand, following the upper echelon theory Adams et al. (2023) suggest a positive influence of the board on the company's innovation performance. As a result of cross-country examinations of a large dataset of 45 countries, increased female participation on boards is positively associated with increased number of patents and corporate innovation (Griffin et al., 2021). Ritter-Hayashi et al. (2019) notes that research focusing on innovation and gender diversity was not carried out in developing countries, and thus expands it to 15 countries. A female representation in board has been proved to increase the chances of firms to implement innovation by 2.19%. Gender diversity in the workforce also has a positive impact on innovation, however the moderation effect of gender equality is found to be insignificant

(Ritter-Hayashi et al., 2019). It is shown that as long as innovation is valued in the company, inclusion of women in top management brings better outcomes in firm performance (Dezsö & Ross, 2012). Similarly, Chen et al. (2018) states that increased female board composition has better monitoring abilities and results in a positive impact on the R&D expenditure and innovation performance of companies that focus on creativity and innovation. Misinterpretation might be caused by self-selection bias, where women initially choose companies with high probability to innovate.

According to Lee and Chung (2022), having more women in top management positions has an influence on the quantity and the impact of innovation, so that although the quantity of innovation increases, the impact of them decreases. This is explained by the risk averse behaviour of women, who take less risky projects, based on the past experience, and have narrow innovation sources. Similarly, a study by Yarram and Adapa (2022) find that having gender diverse corporate boards is linked with lower business risk in the context of Australian firms in ASX 300 Index. This can imply that gender diversity promotes a balanced approach to risk, which prevents businesses from engaging in excessive risk-taking.

Tonoyan and Boudreaux (2023) consider the case of gender diversity in ownership and find out that gender diverse managers in 29 European and Asian emerging economies can significantly affect the propensity to innovate. Additionally, if innovation is taken as an interactive process and the research is scaled to all employee level, the influence of gender diversity on possibility of practicing innovation will be higher (Østergaard et al., 2011).
On the other hand, Almor et al. (2022) explores the relationship between gender diversity on boards and R&D investments, and finds a duality in results in a cross-country analysis: while an increased gender diversity has a significant negative effect on R&D, it also positively influences the organizational behavior, which increases the R&D investments over time.

Drawing upon the critical mass theory as a framework and utilizing panel dataset from Germany, Joecks et al. (2013) observe that there is a U-shaped relationship between gender diversity and firm performance. More specifically, after adjusting for reverse causality, authors conclude that until the 30% threshold is achieved, the increased share of women on boards have an adverse impact on performance. Once companies attain a critical mass of 30% female representation, firms experience better results compared with all-male management boards.

This finding may serve as one of the explanations for the contradictory results observed on previous studies, indicating that it might be plausible that papers with negative evidence focused on companies with female participation lower than a particular threshold. We will evaluate this theory through empirical analysis.

1.2.3 Gender diversity in Spanish firms

The increasing female participation in the workforce of Spain attracted researchers to empirically study the possible outcomes of this change (Teruel & Segarra-Blasco, 2017).
Despite the introduction of quotas and progress made in the labor market, women continue to be underrepresented in Spanish firms. Following the "2007 Ley de Igualdad" law, public companies were supposed to meet the 40% quota by 2015 (Comi et al., 2020). However, gender quotas, which were formally recommended rather than obligated, did not significantly increase the representation of women in corporate boards. More specifically, according to the gender gap reports by World Economic Forum, the gender gap index changed from 0.7444 in 2007 to 0.788 in 2022 (World Economic Forum, 2007; World Economic Forum, 2022). The score ranges from 0 to 1, with a value closer to 1 meaning greater gender equality. Maida and Weber (2019) find similar results in case of Italian Board Reform, consistent with the outcome of a law in Norway, that while the reform led to an increase in the share of women in corporate positions, there is no indication of spillover effect for other high positions.

Still, these studies mostly focus on explaining the effect of gender diversity on the firm's performance and thus it raised our interest to analyse its impact on the innovation indicators. Few articles observing this relationship using the PITEC data, which is also used in our study, have following conclusions. The outcome of studies by Díaz-García et al. (2013) show that gender diversity in R&D departments is positively correlated with radical innovation, however it does not foster incremental innovations. Sastre (2015) centers the research on manufacturing sector and finds that the relationship between gender diversity and innovation outputs have an inverted U shape. The author also shows that the effect of gender diversity differs depending on a sector, having the most significant impact on product innovation, then on organisational innovation, followed by the process innovation (Sastre, 2015).

Teruel and Segarra-Blasco (2017) have an ambiguous outcome: gender diversity in R&D teams has different effect depending on the type of patent, a negative effect on the Spanish patents and a positive effect on internationally registered patents. Moreover, the firm size and age are also significant contributing factors with older and smaller companies having more capacity to patent. More recent study by Becker and Vahter (2022) finds positive effect of the increase in the share of women in R&D department on the probability of pursuing an eco-innovation strategy.

As the existing papers on the PITEC dataset focus on the change within the R&D teams, we'll consider the relationship between gender diversity in the total workforce and the innovation outputs.

1.2.4 Empirical results of open innovation

While literature on the relationship between gender diversity and innovation is limited, the studies on implications of increased gender diversity on open innovation are scarce (Wikhamn & Knights, 2013). Laursen and Salter (2006) were the first to empirically investigate the relationship between the open search strategies and the innovative performance of firms in the United Kingdom. Authors introduced the concept of breadth and depth as factors of the companies' openness to external sources and have found strong evidence of a positive relationship. Firms that draw knowledge from external sources are more likely to see positive change in their innovation performance. However, there might also exist the possibility of over-searching, leading to a curvilinear relationship due to high cost of innovation search (Laursen & Salter, 2006).

In the context of Estonian firms, Masso and Vahter (2020) observe that newly hired women get more benefits working at more open innovative companies, as they are associated with feminine characteristics. However, the wages of women do not match those for men in this scenario too. When examining the factors of gender wage gap, authors identify that conventional determinants do not completely explain the wage gap and that approximately 35% of the gap in Estonian firms can be attributed for differences in firm productivity (Masso and Vahter, 2020). This is mainly explained by the bargaining power difference between men and women, a higher flexibility of men, and a distinct approach to risk-taking. In case of Estonian companies during 2006 to 2018, Masso et al. (2022) find that firm-level factors account for up to 40% of gender wage gap. The results suggest that men tend to sort themselves into high-productivity companies paying higher wages, while women tend to negotiate lower wages within organisations.

The findings of empirical studies on the complementarities between internal and external innovation activities are still unclear. Complementarity refers to a degree to which addition of one activity leads to an increase in the marginal returns of the other. Love et al. (2014) tested the existence of dynamic complementarities of internal R&D and external knowledge sources on the example of panel data from Ireland. The authors did not observe a significant trend of companies adopting more open innovation strategies. However, they find evidence that utilizing both external and internal knowledge sources together is linked with a rise in innovation outputs in the dynamic context. In the case of manufacturing industry in Belgium, Cassiman and Veugelers (2006) observe that companies that focus solely on one innovation activity, such as internal R&D or external technology sourcing, introduce less innovations than companies that engage in open innovation combining these two activities. This suggests that there is complementarity between the internal R&D and acquisition of external knowledge activities.

A recent paper by Adams et al. (2023) focuses on exploring the influence of female board members on open innovation in emerging markets with focus on BRICS countries. Authors note that women possess behavioral characteristics necessary for successful coupled open innovation, and show that gender diverse composition of corporate boards can have a positive impact on the adoption and maintenance of open innovation strategies. However, it should be noted that the extent of impact that board members have on a firm's performance may vary depending on the surrounding business environment (Adams et al., 2023), as Post and Byron (2015) reveal in their meta-analysis that the correlation between female board composition and firm performance is negative with lower gender parity and positive, otherwise.

To sum up, the empirical evidence from prior investigations produced mixed results regarding the association between gender diversity and both innovation outputs and inputs. Moreover, these studies primarily examined the representation of women on R&D teams or corporate boards, leaving uncertainty regarding the relationship between gender diversity across the entire workforce and innovation performance. An examination of past studies through meta-analysis suggests that the variances in findings can be partially accounted for by the cultural and regulatory context in which the studies were conducted. These findings reveal the significance of taking into account the wider social context surrounding organizations when observing the correlation between gender diversity and firm performance.

2. Data and Methodology

2.1 Data

The empirical analysis will be conducted using Spanish Technological Innovation Panel (PITEC) dataset at the firm level, collected through collaboration of the Spanish National Statistics Institute and the Spanish Foundation for Science and Technology. The data was gathered between 2003 and 2016 for more than 12,000 companies on an annual basis through "Innovation in Companies" surveys (Teruel & Segarra-Blasco, 2017).

This dataset has its strengths and weaknesses. Firstly, the survey follows the classifications in Oslo Manual instructions (OECD&Eurostat, 2005), is oriented towards firms that demonstrate capacity for innovation, and is a Spanish contribution to European Community Innovation Survey (CIS) (Rodriguez et al., 2017; Teruel & Segarra-Blasco, 2017). Secondly, The PITEC dataset consists of more than 170,000 firm-year observations. The panel data allows us to observe long term relationships between variables, details of innovation behavior within the same firms over years (Segarra & Teruel, 2014). The survey includes rich variables to measure innovation activities of the companies, as well as other different firm characteristics.

Based on the methodology notes along with past studies, the PITEC data comprises of companies involved in innovation activities, indicating its representation of manufacturing and service firms (Bianchini et al., 2018). However, as highlighted by Segarra and Teruel (2014), the data might not be fully representative of the entire population. Duch-Brown et al. (2016) attribute the overrepresentation of innovative, productive, and larger companies compared to CIS surveys to the design of the survey. In addition, we do not observe exit and entrance of firms, and focus on surviving firms.

We limit the data from 2004 to 2016, when all necessary variables are available. Initially, the dataset has 171,511 observations in 2003-2016 period. Since the year 2003 does not have information related to knowledge sources, after removing instances we have 164,228 observations left. Considering that only firms with "no incident" as their incident indicator have filled in the questionnaire, we omit the remaining companies who didn't respond to any survey questions. Dropping 35,656 missing information reduces the data to 128,572 observations. Exploring missing variables, descriptive statistics in **Table 2** shows the number of observations after removing NAs for each variable. Thus, the questions about the cooperation and external knowledge sources were skipped the most, resulting in the lowest response level.

It is important to note that not all questions require answers, i.e., if a company didn't introduce an innovation in the given period, the respondent should skip the questions related to innovation and move on to other sections. Hence, based on the structure of a survey we decide to take the answer as 0 or NA. It should also be taken into account that answers correspond to different time periods with the main year being the year prior the data collection. Some variables such as the number of employees, group ownership, R&D expenses, financing are related with one-year period, while the variables regarding the innovation such as cooperation, knowledge sources, radical and incremental innovations refer to three-year period. The main activity of companies is also an important factor. The firms in our sample operate in different industries such as manufacture, service, construction, and agriculture. Our study will consider the more detailed examination of the manufacturing and service industries. As regards the software tools, STATA and R be used for data processing.

2.2 Definitions of variables

2.2.1 Explanatory variables

Increase in the share of women is chosen as the main explanatory variable. This is a binary variable, taking value "1" when the change of share of women in the current year as compared with the previous year is positive, and "0" otherwise. We introduced a 10% threshold looking at the distribution of change in share of women in **Figure 1.** Therefore, we generate three treatment groups: more than 0%, between 0 and 10% and more than 10% increase in share of women.

As regards gender diversity, there are different ways to measure it: Blau index, Shannon–Weaver entropy index, Herfindahl-Hirschman index, among others. As Blau index is the standard measure of diversity in the binary treatment, it will be utilized in our empirical analysis. As shown in equation **(1),** *B* represents the value of Blau index, *p* refers to the proportion of members in the categories from *i* to *k* (Teruel & Segarra-Blasco, 2017; Xie et al.). In our case, we consider 2 categories for *i:* men and women. The output value is continuous and can range between 0 and 0.5, with the former referring to the gender homogeneity and the latter to the gender diversity.

$$B = 1 - \sum_{i=1}^{k} p_i^2 \quad (1)$$

Based on this Blau index, we create a binary variable for an increase in blau index. The **Figure 2** exhibits how the change in Blau index is distributed in the dataset. By considering it, we have decided on 0.1 threshold for creating treatment groups. Three treatment groups are formulated based on more than 0, between 0 and 0.1, more than 0.1 increase in Blau index.

2.2.2 Dependent variables

For the indicator of innovation, we have several variables. Product innovation is a binary variable, coded as 1, when a company introduced a goods or services related innovation in the period, and 0, otherwise. Similarly, we have a process innovation variable.

We create a variable of open innovation, based on the combination of two variables: the use of internal R&D and external knowledge sourcing (Love et al., 2014). According to the survey structure we used the question "Did the company introduce the innovated goods or services during the three year period?" In case of positive response, firms are asked to indicate the sources from which the information was taken to complete projects and a degree of importance. More specifically, four types of information sources were identified: internal sources, sources from suppliers and competitors, institutional sources and other sources. As regards the degree of importance, we first group the high and medium together, and low and not used together to create a binary variable.

The variable cooperation is a binary variable taking "1" when a company collaborates with other companies in innovation activities, and "0", otherwise. As alternative indicators for open innovation we consider the variables for external knowledge sourcing and cooperation separately.

2.2.3 Control variables

The selection of control variables is primarily justified by past studies, specifically those that explored the relationship between gender diversity and innovation. Díaz-García et al. (2013) considered factors such as company size, R&D team size, external R&D, and group membership as covariates. Becker and Vahter (2022) additionally includes training, barriers of innovation, turnover, foreign ownership, and the status of export.

Control variables that are likely to affect the firm innovation and increase in the share of women in our study are as follows. The number of total employees relates to the size of a firm and it was linked to innovation capabilities in literature (Díaz-García et al., 2013). Additionally, we consider the number of employees in the R&D teams and a percentage of women in the total workforce and in the R&D teams. Team size is also considered to be a key indicator of team performance (Díaz-García et al. 2013). These are initially provided as continuous variables, but we calculated the logarithms to avoid the data skewness.

We create the dummy variables for Internal R&D and External R&D, referring to the activities of technological innovation. They take value 1, when the company performed these activities and 0 if it didn't. We use the variables for incremental and radical innovation, where the difference between them lies in whether the innovation is introduced for the firm or for the market. These dummy variables, which take value 1 and 0, indicate whether or not the company has introduced this type of innovation in the survey period.
Next, group membership is a binary variable, coded 1 if the company is a part of a group of companies and 0 otherwise. Similarly, the foreign ownership control variable was created. Neither variable is opposite from open innovation variable and shows the combination of no R&D and external knowledge sourcing (Love et al. 2014).

We turn the continuous variables for public funds and export level to binary variables as well. As a measure of barriers to innovation, we include the lack of qualified staff. As for technological innovation activities, we include acquisition of machinery, training, market research, design as binary variables if company performed any of these activities. Based on the amount of turnover, we calculate sales per person and take a logarithm of that. We also consider the sector and year dummies to obtain more accurate results.

2.3 Descriptive Statistics

Table 2 displays the mean, standard deviation of the variables and a number of observations used in the analysis. The proportion of women constitutes the 32.5% during the study period. The share of women in the R&D departments is less and equals 12.8%. We find that Blau index is equal to 0.313 in the total workforce and 0.115 in the R&D teams respectively in 128,572 observations from our sample. This indicates that women are underrepresented in both total workforce and R&D teams. If 56.4% of firms on average had increase in the share of women in total workforce, only 13.2% of firms had an increase in the proportion of

women compared with the prior year in R&D. This can show the persistency of the low representation of women in R&D. Additionally, increase in Blau index was observed in 55.7% of companies.
Removing the data for internal sources, we obtain the mean number of external sources of 2.5 with a standard deviation of 2.86. On average, 81.2% of firms use external knowledge sourcing to implement innovation activities. As open innovation refers to utilizing both internal R&D and external source of information, 59% of companies engage in open innovation.

Table 3 also presents the average values for manufacturing and service industry firms. In total, a bigger share of the firms belongs to the manufacturing industry (50.46%). Firms operating in the service industry (41.79% of total) are more gender-diverse than manufacturing ones as Blau index indicates. However, on average, manufacturing firms are involved in innovation practices, such as product, process, open, radical, incremental innovation, more than service firms.

As the PITEC dataset doesn't include data on the proportion of women working on corporate boards, it's not feasible to examine the impact of 40% gender quotas and potential spillover effects. **Figure 3** indicates that the average percentage of women in the overall workforce between 2004-2016 remained within the range of 31% to 34%. **Figure 4** demonstrates **the** share of treatment groups observations in total observations. We can see that number of observations for 10% threshold in increased share of women variable corresponds to 0.1 positive change in Blau index variable.

Table 2

Descriptive Statistics

Variables	*Observations*	*Mean*	*Standard deviation*
Number of total employees (ln)	128572	4.139	1.728
Number of total employees in R&D (ln)	128572	0.912	1.256
Share of women employees	128572	0.325	0.250
Share of women employees in R&D	128572	0.128	0.232
Increase in share of women	75298	0.564	0.496
Increase in share of women in R&D	97783	0.132	0.359
Blau index	128572	0.313	0.158
Blau index in R&D	128572	0.115	0.185
Increase in Blau index	75119	0.557	0.497
Group membership	128572	0.410	0.492
Internal R&D	128572	0.479	0.500
External R&D	128572	0.226	0.418
External knowledge sourcing	94286	0.812	0.391
Cooperation	94287	0.369	0.483
Incremental innovation	128572	0.354	0.478
Radical innovation	128572	0.268	0.443
Product innovation	128572	0.463	0.499
Process innovation	128572	0.473	0.499
Open innovation	94286	0.592	0.491
Neither	94286	0.128	0.334
Foreign ownership	128572	0.142	0.349
Public funding	116420	0.147	0.354
Exports	128572	0.417	0.493
Lack of employees	128572	0.369	0.483
Sales per person	128572	11.681	1.165
Number of external knowledge sources	128572	2.528	2.862
Machine	128572	0.170	0.376
External knowledge acquisition	128572	0.029	0.167
Trainings	128572	0.130	0.336
Market research	128572	0.173	0.378
Design	128572	0.072	0.258

Note. NA values are excluded.
Source*:* authors' calculations from PITEC

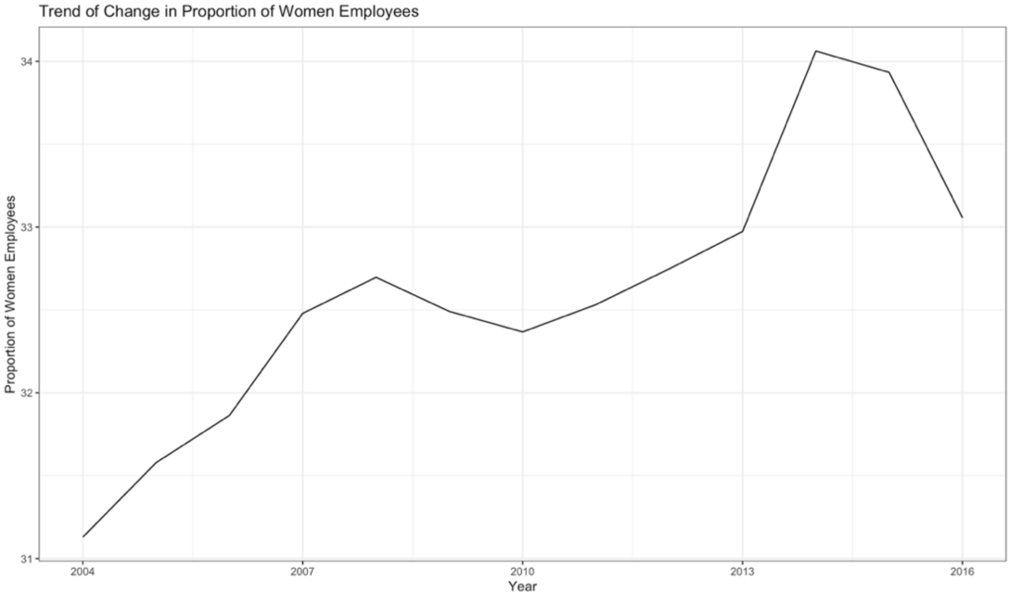

Figure 3. The average proportion of women employees in 2004-2016
Source: authors' calculations from PITEC

Table 4 presents the correlations between main variables. The result reveals no significant association between increase in share of women and open innovation. While increase in share of women employees in total workforce doesn't have remarkable correlation with other variables, increase in share of women in R&D is positively related with external R&D, external knowledge sourcing and open innovation at 5% significance level. Open innovation also positively correlates with proportion of women in R&D, Blau index in R&D, cooperation, radical innovation. While there is correlation between several variables, most of their sizes are less than 0.5. Considering non-evident correlations between other variables, we proceed with further analysis to see the difference in means when companies have and don't have an increase in the share of women in the total workforce.

Table 4

Correlation Table

	Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Share of women employees in R&D	1														
2	Increase in share of women	-0.02	1													
3	Increase in share of women in R&D	0.44***	0.15	1												
4	Blau index in R&D	0.66***	0.04	0.59***	1											
5	Increase in Blau index	-0.02	0.99**	0.15	0.04	1										
6	Internal R&D	0.5***	0	0.41***	0.57**	0	1									
7	External R&D	0.18*	0.02	0.23*	0.26**	0.02	0.28**	1								
8	External knowledge sourcing	0.17*	0.02	0.15*	0.21**	0.02	0.31***	0.18*	1							
9	Cooperation	0.16	0.02	0.21*	0.25*	0.02	0.24**	0.31***	0.24**	1						
10	Incremental innovation	0.09	0.03	0.09	0.12	0.03	0.17	0.08	0.12	0.09	1					
11	Radical innovation	0.1	0.02	0.15	0.18	0.02	0.25*	0.14	0.16*	0.18	0.15*	1				
12	Product innovation	0.1	0.03	0.14	0.16	0.03	0.24	0.1	0.17	0.15	0.74***	0.57***	1			
13	Process innovation	-0.01	0.04	0.04	0.01	0.04	-0.05	0.04	0.09	0.1	0.04	0.04	0.01	1		
14	Open innovation	0.46***	0.01	0.39***	0.54***	0.01	0.9***	0.29**	0.52***	0.28**	0.18	0.27*	0.26	-0.01	1	
15	Number of external sources	0.24**	0.01	0.23**	0.31***	0.01	0.35***	0.26**	0.57***	0.33***	0.15	0.2*	0.19	0.07	0.47***	1

Notes. (*) represents a significance level of 0.05 (5%), (**) represents a significance level of 0.01 (1%), (***) represent a significance level of 0.001 (0.1%).
Source: authors' calculations from PITEC

The results of comparison of means are provided in **Tables 5-7**. By keeping the control group unchanged, we perform the t-test for 3 different treatment groups: more than 0% change in share of women, between 0 and 10% increase in share women, more than 10% increase in share of women. Specifically, 0 in the table stands for no increase in the share of women and 1 refers to the increase. Results of **Table 5** shows that more firms, 42,472, had a positive change in the proportion of women as compared with 32,826 companies who didn't have any alteration. For the next treatment groups, the number of observations declines to 32,448 and 19,658 as shown in **Table 6 and 7.** Organizations who had an increase in the female representation have, on average, engaged more in innovation activities (incremental, radical, product, process innovation, cooperation) at <0.001 p-value for all three treatment groups. However, the proportion of firms utilizing open innovation isn't statistically different for firms that have any positive change in the proportion of women in their total workforce. Moreover, **Table 6** suggests that for firms with increase between 0 and 10% incorporating open innovation increases compared to the ones with no change in the share of women. Opposite results are found for the third treatment group with an increase of more than 10% in **Table 7**. Utilizing open innovation activities decreases when companies increase their share of women more than 10%. These results suggest a curvilinear relationship between open innovation and increase in share of women.

When carrying out the means test based on increase in Blau index variable, we create similar treatment groups: companies with a positive change in their Blau index, with an increase in Blau between 0 and 0.1, increase more than 0.1 in **Tables 8-10**. Firms with a positive change in gender diversity tend to have more incentive to participate in innovation related activities, as can be seen in **Table 8.** However, for open innovation, there is no significant change in the case of the first treatment group. The share of companies that engaged open innovation to their working process is higher for companies that increased their Blau index between 0 and 0.1 compared to last year from **Table 9**. More than 0.1 increase in gender diversity indicator is associated with decreasing level of open innovation, as shown in **Table 10**. Thus, the result is in line with previous findings implying the existence of inverted U-shaped relationship for open innovation.

We also checked the difference in means test for the decrease in share of women and decrease in gender diversity indicator. **Table 12** presents the result comparing the mean values of firms with no change and decrease in share of women. When it comes to open innovation, there is no significant difference between the two groups. However, on average, smaller number of firms apply open innovation in the group with decrease in Blau index as shown in **Table 11.**

Table 6

Mean t-test. The comparison between companies with no change and an increase between 0 and 10% in increase in share of women

	Increase between 0 and 10%			
	Group 0 Mean (SD)	*Group 1 Mean (SD)*	*Difference*	*Significance*
Group membership	0.38 (0.49)	0.47 (0.50)	-0.09	***
Internal R&D	0.46 (0.50)	0.50 (0.50)	-0.03	***
External R&D	0.21 (0.40)	0.24 (0.43)	-0.04	***
External knowledge sourcing	0.83 (0.37)	0.86 (0.35)	-0.02	***
Cooperation	0.38 (0.48)	0.41 (0.49)	-0.03	***
Incremental innovation	0.34 (0.47)	0.38 (0.49)	-0.04	***
Radical innovation	0.26 (0.44)	0.29 (0.45)	-0.03	***
Product innovation	0.44 (0.50)	0.49 (0.50)	-0.05	***
Process innovation	0.45 (0.50)	0.51 (0.50)	-0.07	***
Open innovation	0.61 (0.49)	0.63 (0.48)	-0.02	***
Neither	0.11 (0.31)	0.10 (0.30)	0.01	***
Foreign ownership	0.14 (0.35)	0.17 (0.37)	-0.03	***
Public funding	0.14 90.35)	0.16 (0.37)	-0.02	***
Exports	0.40 (0.49)	0.45 (0.49)	-0.06	***
Lack of employees	0.37 (0.48)	0.37 (0.48)	0.00	NS
Sales per person	11.68 (1.20)	11.79 (1.09)	-0.11	***
Number of external knowledge sources	2.47 (2.86)	2.75 (2.90)	-0.28	***
Machine	0.15 (0.35)	0.17 (0.38)	-0.03	***
External knowledge acquisition	0.02 (0.14)	0.03 (0.16)	-0.01	***
Trainings	0.10 (0.30)	0.12 (0.33)	-0.02	***
Market research	0.16 (0.37)	0.18 (0.38)	-0.02	***
Design	0.06 (0.23)	0.06 (0.24)	-0.01	***
Number of observations	32,826	32,448		

Notes. Welch Two Sample t-test; NA values are excluded. Standard errors are presented in parentheses.
Source: authors' calculations from PITEC

Table 7

Mean t-test. The comparison between companies with no change and an increase more than 10% increase in share of women

	Increase more than 10%			
	Group 0 *Mean (SD)*	*Group 1* *Mean (SD)*	*Difference*	*Significance*
Group membership	0.38 (0.49)	0.45 (0.50)	-0.07	***
Internal R&D	0.46 (0.50)	0.46 (0.50)	0.01	***
External R&D	0.21 (0.40)	0.22 (0.42)	-0.02	***
External knowledge sourcing	0.83 (0.37)	0.84 (0.37)	0.00	NS
Cooperation	0.38 (0.48)	0.40 (0.49)	-0.02	***
Incremental innovation	0.34 (0.47)	0.35 (0.48)	-0.01	***
Radical innovation	0.26 (0.44)	0.27 (0.44)	-0.01	***
Product innovation	0.44 (0.50)	0.46 (0.50)	-0.02	***
Process innovation	0.45 (0.50)	0.48 (0.50)	-0.03	***
Open innovation	0.61 (0.49)	0.61 (0.49)	0.01	**
Neither	0.11 (0.31)	0.12 (0.32)	-0.01	**
Foreign ownership	0.14 (0.35)	0.16 (0.36)	-0.02	***
Public funding	0.14 (0.35)	0.15 (0.35)	-0.01	***
Exports	0.40 (0.49)	0.40 (0.49)	0.00	NS
Lack of employees	0.37 (0.48)	0.37 (0.48)	0.01	NS
Sales per person	11.68 (1.20)	11.71 (1.19)	-0.03	**
Number of external knowledge sources	2.47 (2.86)	2.57 (2.92)	-0.10	***
Machine	0.15 (0.35)	0.16 (0.37)	-0.02	***
External knowledge acquisition	0.02 (0.14)	0.02 (0.15)	0.00	**
Trainings	0.10 (0.30)	0.11 (0.32)	-0.01	***
Market research	0.16 (0.37)	0.17 (0.37)	-0.01	*
Design	0.06 (0.23)	0.06 90.24)	0.00	*
Number of observations	32,826	19,658		

Notes. Welch Two Sample t-test; NA values are excluded. Standard errors are presented in parentheses.

Source: authors' calculations from PITEC

Table 9

Mean t-test. The comparison between companies with no change and an increase between 0 and 0.1 in their Blau index

	Increase between 0 and 0.1			
	Group 0 Mean (SD)	*Group 1 Mean (SD)*	*Difference*	*Significance*
Group membership	0.38 (0.49)	0.47 (0.50)	-0.09	***
Internal R&D	0.46 (0.50)	0.50 (0.50)	-0.04	***
External R&D	0.21 (0.40)	0.24 (0.43)	-0.04	***
External knowledge sourcing	0.83 (0.37)	0.86 (0.36)	-0.02	***
Cooperation	0.38 (0.48)	0.41 (0.49)	-0.03	***
Incremental innovation	0.34 (0.47)	0.38 (0.49)	-0.04	***
Radical innovation	0.25 (0.44)	0.29 (0.45)	-0.03	***
Product innovation	0.44 (0.50)	0.49 (0.50)	-0.05	***
Process innovation	0.45 (0.50)	0.52 (0.50)	-0.07	***
Open innovation	0.61 (0.49)	0.63 (0.48)	-0.02	***
Neither	0.11 (0.31)	0.10 (0.30)	0.01	***
Foreign ownership	0.14 (0.35)	0.17 (0.37)	-0.03	***
Public funding	0.14 (0.35)	0.16 (0.37)	-0.02	***
Exports	0.40 (0.49)	0.45 (0.50)	-0.05	***
Lack of employees	0.37 (0.48)	0.37 (0.48)	0.00	NS
Sales per person	11.68 (1.20)	11.78 (1.11)	-0.10	***
Number of external knowledge sources	2.46 (2.86)	2.77 (2.92)	-0.30	***
Machine	0.15 (0.35)	0.17 (0.38)	-0.02	***
External knowledge acquisition	0.02 (0.14)	0.02 (0.15)	0.00	***
Trainings	0.10 (0.30)	0.13 (0.33)	-0.02	***
Market research	0.16 (0.37)	0.18 (0.38)	-0.02	***
Design	0.06 (0.23)	0.06 (0.24)	-0.01	***
Number of observations	33,270	33,936		

Notes. Welch Two Sample t-test; NA values are excluded. Standard errors are presented in parentheses.
Source: authors' calculations from PITEC

Table 10

Mean t-test. The comparison between companies with no change and an increase more than 0.1 in their Blau index

	Increase more than 0.1			
	Group 0 Mean (SD)	*Group 1 Mean (SD)*	*Difference*	*Significance*
Group membership	0.38 (0.49)	0.35 (0.48)	0.03	***
Internal R&D	0.46 (0.50)	0.39 (0.49)	0.07	***
External R&D	0.21 (0.40)	0.18 (0.38)	0.03	***
External knowledge sourcing	0.83 (0.37)	0.80 (0.40)	0.03	***
Cooperation	0.38 (0.48)	0.34 (0.47)	0.04	***
Incremental innovation	0.34 (0.47)	0.31 (0.46)	0.02	***
Radical innovation	0.25 (0.44)	0.23 (0.42)	0.02	***
Product innovation	0.44 (0.50)	0.42 (0.49)	0.03	***
Process innovation	0.45 (0.50)	0.43 (0.49)	0.02	***
Open innovation	0.61 (0.49)	0.53 (0.50)	0.07	***
Neither	0.11 (0.31)	0.14 (0.35)	-0.03	***
Foreign ownership	0.14 (0.35)	0.11 (0.31)	0.03	***
Public funding	0.14 (0.35)	0.12 (0.33)	0.02	***
Exports	0.40 (0.49)	0.33 (0.47)	0.07	***
Lack of employees	0.37 (0.48)	0.37 (0.48)	-0.01	NS
Sales per person	11.68 (1.20)	11.58(1.23)	0.10	***
Number of external knowledge sources	2.46 (2.86)	2.22 (2.77)	0.25	***
Machine	0.15 (0.35)	0.15 (0.36)	0.00	NS
External knowledge acquisition	0.02 (0.14)	0.02 (0.13)	0.00	NS
Trainings	0.10 (0.30)	0.08 (0.28)	0.02	***
Market research	0.16 (0.37)	0.14 (0.35)	0.02	***
Design	0.06 (0.23)	0.05 (0.22)	0.00	*
Number of observations	33,270	7,913		

Notes. Welch Two Sample t-test; NA values are excluded. Standard errors are presented in parentheses.
Source: authors' calculations from PITEC

Table 11

Mean t-test. The comparison between companies with no change and a negative change in their Blau index

	Decrease in Blau index			
	Group 0 *Mean (SD)*	*Group 1* *Mean (SD)*	*Difference*	*Significance*
Group membership	0.38 (0.49)	0.43 (0.49)	-0.05	***
Internal R&D	0.46 (0.50)	0.46 (0.50)	0.01	NS
External R&D	0.21 (0.40)	0.22 (0.41)	-0.01	***
External knowledge sourcing	0.83 (0.37)	0.84 (0.37)	0.00	NS
Cooperation	0.38 (0.48)	0.38 (0.4)	0.00	NS
Incremental innovation	0.34 (0.47)	0.36 (0.48)	-0.02	***
Radical innovation	0.25 (0.44)	0.27 (0.45)	-0.02	***
Product innovation	0.44 (0.50)	0.47 (0.50)	-0.03	***
Process innovation	0.45 (0.50)	0.48 (0.50)	-0.03	***
Open innovation	0.61 (0.49)	0.59 (0.49)	0.02	***
Neither	0.11 (0.31)	0.11 (0.32)	-0.01	**
Foreign ownership	0.14 (0.35)	0.15 (0.35)	-0.01	***
Public funding	0.14 (0.35)	0.15 (0.36)	-0.01	***
Exports	0.40 (0.49)	0.41 (0.49)	-0.01	***
Lack of employees	0.37 (0.48)	0.37 (0.48)	0.00	NS
Sales per person	11.68 (1.20)	11.67 (1.13)	0.01	NS
Number of external knowledge sources	2.46 (2.86)	2.58 (2.90)	-0.12	***
Machine	0.15 (0.35)	0.15 (0.36)	-0.01	***
External knowledge acquisition	0.02 (0.14)	0.02 (0.14)	0.00	NS
Trainings	0.10 (0.30)	0.11 (0.31)	-0.01	***
Market research	0.16 (0.37)	0.16 (0.37)	-0.01	*
Design	0.06 (0.23)	0.06 (0.23)	0.00	NS
Number of observations	33,270	38,211		

Notes. Welch Two Sample t-test; NA values are excluded. Standard errors are presented in parentheses.
Source: authors' calculations from PITEC

Table 12

Mean t-test. The comparison between companies with no change and a negative change in their share of women

	Decrease in share of women			
	Group 0 *Mean (SD)*	*Group 1* *Mean (SD)*	*Difference*	*Significance*
Group membership	0.38 (0.49)	0.45 (0.50)	-0.07	***
Internal R&D	0.46 (0.50)	0.47 (0.50)	-0.01	NS
External R&D	0.21 (0.40)	0.23 (0.42)	-0.02	***
External knowledge sourcing	0.83 (0.37)	0.85 (0.36)	-0.01	***
Cooperation	0.38 (0.48)	0.39 (0.49)	-0.01	***
Incremental innovation	0.34 (0.47)	0.37 (0.48)	-0.03	***
Radical innovation	0.25 (0.44)	0.28 (0.45)	-0.02	***
Product innovation	0.44 (0.50)	0.48 (0.50)	-0.04	***
Process innovation	0.45 (0.50)	0.50 (0.50)	-0.05	***
Open innovation	0.61 (0.49)	0.60 (0.49)	0.01	NS
Neither	0.11 (0.31)	0.11 (0.31)	0.00	NS
Foreign ownership	0.14 (0.35)	0.16 (0.37)	-0.02	***
Public funding	0.14 (0.35)	0.15 (0.36)	-0.01	***
Exports	0.40 (0.49)	0.43 (0.49)	-0.03	***
Lack of employees	0.37 (0.48)	0.37 (0.48)	0.00	NS
Sales per person	11.68 (1.20)	11.72 (1.12)	-0.04	***
Number of external knowledge sources	2.46 (2.86)	2.64 (2.91)	-0.17	***
Machine	0.15 (0.35)	0.16 (0.37)	-0.01	***
External knowledge acquisition	0.02 (0.14)	0.02 (0.14)	0.00	NS
Trainings	0.10 (0.30)	0.12 (0.32)	-0.01	***
Market research	0.16 (0.37)	0.17 (0.38)	-0.01	**
Design	0.06 (0.23)	0.06 (0.23)	0.00	**
Number of observations	33,270	47,666		

Notes. Welch Two Sample t-test; NA values are excluded. Standard errors are presented in parentheses.
Source: authors' calculations from PITEC

2.4 Methodology

2.4.1 Endogeneity and reverse causality

It is important to address the endogeneity issue when assessing the relationship between an increased share of women and firm's introduction of innovation. It is challenging to determine whether having a gender diverse workforce contributes to an increase in firm's innovation outputs or whether highly innovative companies attract a more gender diverse employees. Endogeneity problem occurs when the treatment variable, increase in the share of women, and the error term are correlated because there is a probability of reverse causality and omitted variables. Hence, the inclusion of women might not be caused by external factors, but other factors like self-selection process and omitted variables bias (Yang et al., 2019). Thus, we will focus on discussing the association between variables.

The empirical analysis will be conducted using the propensity score matching technique (PSM) (Rosenbaum & Rubin, 1983). This method helps us to take into account number of firm characteristics like size, sector, group ownership, and include performances of the corresponding and past years. Moreover, PSM method was effectively utilized on papers using the same dataset (Becker & Vahter, 2022). Although matching is widely used to estimate the causal effect, PSM has limitations and due to factors, such as unobserved fixed characteristics of firms, PSM can account for the endogeneity concern only up to some extent (Sainani, 2012).

2.4.2 Propensity score matching

The steps of PSM method are as follows (Caliendo & Kopeinig, 2008).

The first step is a propensity score estimation. A propensity score estimation requires the selection of a model and the suitable variables. We use probit regression model (2) since we have the case of a binary treatment. It is important to include covariates, as omitting them can affect the bias, while including non-significant variables might affect the variance. Hence, for control variables we consider those variables that can affect both the gender diversity at workforce and innovation performance, listed in section 2.2.3. The selection of variables depends on the economic theory and prior empirical studies. In equation (2), y_{it} stands for an increase in the share of women at t, , x_k denotes the control variables lagged by three years, i - the idenfication code of firms, ε is an error term

Probit estimation equation: $y_{it} = \beta_0 + \beta_1 x_{1,i(t-3)} + \beta_2 x_{2,i(t-3)} + .. + \beta_k x_{k,i(t-3)} + \varepsilon$ (2)

Next step is choosing a matching algorithm. There are several PSM estimators, such as Nearest Neighbour, Caliper and Radius, and Stratification and Interval, which show different trade-off between bias and efficiency. We will proceed with Nearest Neighbour Matching, as it is the most common method and matches the control and treated group members, based on the similarity of the propensity score. A drawback of this matching principle is that if the scores of two groups differ significantly, less participants will be matched. However, this can be improved by replacement and oversampling.

The following step is to estimate the average treatment effect on the treated (ATT) of the companies that had an increase in the share of women. Then, it is essential to check common support condition to ensure the comparability of two groups. The common support refers to the range in the propensity score, where the propensity score of control and treatment groups overlaps. In this regard, we use the minima and maxima comparison. More specifically, we remove all observations that are less than the minimum value of one group and greater than the maximum value of another group. After that we check the matching quality and conduct a sensitivity analysis.

ATT equation: $ATT = E(y^1_{t+3} - y^0_{t+3} \mid D_t = 1) = E(y^1_{t+3} \mid D_t = 1) - E(y^0_{t+3} \mid D_t = 1)$ (3)

3. Results and discussion

3.1 Results of PSM

We begin with plotting the propensity scores to check their distribution. **Figure 5** shows that propensity scores of the treatment and control groups overlap sufficiently in the sample, which confirms that the common support condition is fulfilled. All treatment groups in our analysis follow a similar pattern. Therefore, it is feasible to conduct the propensity score matching.

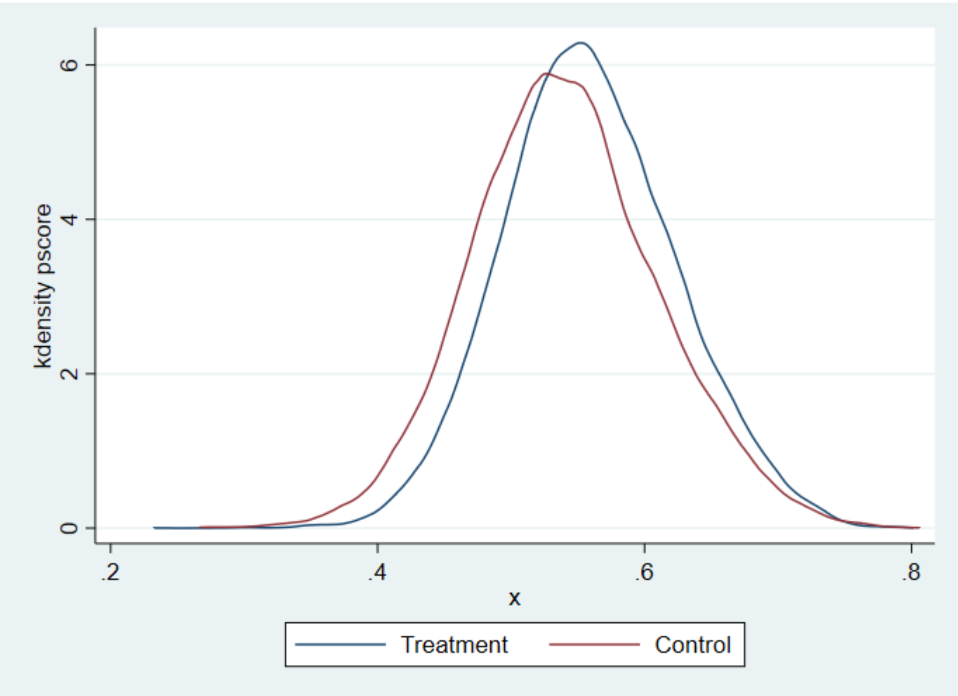

Figure 5. Kernel density estimate of the distribution of the Propensity Score
Source: author's calculations

In **Table 13,** we estimate the probabilities that firm increases share of total women employees utilizing probit regression. The regression is performed for all treatment groups. All control variables have been lagged for t-3 years. We created 4 regression models considering the set of control variables for each outcome. Model 1, 2, 3 and 4 include the control variables of propensity score matching analysis for external knowledge sourcing/cooperation, product innovation, process innovation and open innovation, respectively. When setting the treatment group with more than 0% increase in share of women, looking at statistically significant coefficient estimates shown in **Table 13a**, we can see that firms that are part of a group of companies are more likely to increase share of women in all models. The possibility to increase the proportion of women rises if the company implements staff training. For the second treatment group with an increase between 0% and 10%, a positive relationship is revealed in all models for firms with higher sales productivity per employee at 95% confidence level, as shown in **Table 13b.** There is more likelihood of increasing gender diversity if the firm undertakes exporting activities. Considering the treatment group of more than 10% increase, **Table 13c** reveals that acquisition of machinery and equipment for innovation process is negatively associated with the increase in share of women and it is applicable for all models.

When it comes to gender diversity as treatment variable, the same logical sequence follows for creating regression models of **Table 14**. For models of both treatment group 1 and 2, firms that engage in export operations, have higher sales per staff are more likely to have higher gender representation. However, in the treatment group with more than 0.1 increase in Blau index, inverse relationship exists with carrying out exporting activities and it is valid for all models.

Table 13

Probit regression results for increase in share of women

a) Treatment group 1

	Model 1 Coefficient (Std. error)	Model 2 Coefficient (Std. error)	Model 3 Coefficient (Std. error)	Model 4 Coefficient (Std. error)
Group	0.051*** (0.024)	0.041** (0.021)	0.041** (0.021)	0.052** (0.024)
Share of women	0.172*** (0.053)	0.149*** (0.042)	0.148*** (0.042)	0.169*** (0.053)
Employment	0.023*** (0.008)	0.055*** (0.007)	0.054*** (0.007)	0.022*** (0.008)
Employment in R&D	0.014 (0.013)	-0.007 (0.011)	-0.007 (0.011)	0.014 (0.013)
Internal R&D	-0.102*** (0.035)	-0.011 (0.028)	-0.009 (0.028)	-0.104*** (0.036)
External R&D	0.020 (0.022)	0.017 (0.020)	0.016 (0.020)	0.020 (0.022)
Foreign ownership	-0.032 (0.304)	-0.033 (0.027)	-0.033 (0.027)	-0.039 (0.030)
Cooperation	0.022 (0.023)	0.018 (0.020)	0.017 (0.020)	0.019 (0.023)
External knowledge sourcing	-0.024 (0.035)	-0.014 (0.027)	-0.016 (0.027)	-0.077 (0.049)
Public funding	0.033 (0.026)	0.047* (0.024)	0.047* (0.024)	0.035 (0.027)
Incremental innovation	-0.011 (0.021)	-0.024 (0.028)	-0.004 (0.018)	-0.043 (0.031)

Radical innovation	-0.007 (0.021)	-0.024 (0.024)	-0.009 (0.019)	-0.031 (0.027)
Exports	-0.079 (0.030)	0.042** (0.019)	0.042** (0.019)	0.029 (0.023)
Lack of employees	0.029** (0.023)	0.039** (0.017)	0.039** (0.017)	0.050** (0.020)
Sales per person	0.022* (0.012)	0.028*** (0.010)	0.027*** (0.010)	0.022* (0.012)
Number of external knowledge sources	0.000 (0.004)	0.002 (0.004)	0.002 (0.004)	0.000 (0.005)
Machine	-0.042* (0.023)	-0.127** (0.044)	-0.045** (0.021)	-0.051** (0.024)
External knowledge	-0.076 (0.049)	0.075*** (0.025)	-0.127*** (0.044)	-0.077 (0.049)
Trainings	0.079*** (0.028)	0.075*** (0.025)	0.073*** (0.025)	0.078*** (0.028)
Design	-0.005 (0.034)	0.007 (0.031)	0.005 (0.031)	-0.006 (0.036)
Market research	0.042* (0.024)	0.035 (0.022)	0.038* (0.022)	0.039 (0.025)
Product innovation		0.034 (0.035)		0.055 (0.041)
Process innovation			0.025 (0.020)	0.038 (0.024)
Neither				-0.014 (0.060)
Sector and year dummies	Yes	Yes	Yes	Yes
R^2	0.009	0.012	0.012	0.010
LR-chi^2	217	392.59	393.22	221.40
p>chi^2	0.000	0.000	0.000	0.000
No of obs.	17138	24184	24184	17138

Notes. Model 1, 2, 3 and 4 stands for the regression models using set of control variables for each outcome, namely external knowledge sourcing/cooperation, product/process innovation and open innovation, respectively. Significance levels are ***p < 1%, **p < 5%, *p < 10%.
Source: authors' calculations from PITEC

b) Treatment group 2

	Model 1 Coefficient (Std. error)	Model 2 Coefficient (Std. error)	Model 3 Coefficient (Std. error)	Model 4 Coefficient (Std. error)
Group	0.035 (0.026)	0.022 (0.022)	0.022 (0.022)	0.036 (0.026)
Share of women	0.066 (0.056)	0.040 (0.045)	0.040 (0.045)	0.064 (0.056)
Employment	0.054*** (0.009)	0.093*** (0.007)	0.092*** (0.007)	0.053*** (0.009)
Employment in R&D	0.022* (0.013)	-0.002 (0.012)	-0.002 (0.012)	0.022 (0.013)
Internal R&D	-0.096** (0.037)	-0.000 (0.030)	0.002 (0.030)	-0.098** (0.037)
External R&D	0.017 (0.024)	0.016 (0.021)	0.016 (0.021)	0.017 (0.024)
Foreign ownership	-0.037 (0.317)	-0.031 (0.028)	-0.031 (0.028)	-0.038 (0.317)
Cooperation	0.019 (0.024)	0.015 (0.021)	0.013 (0.021)	0.01 (0.024)
External knowledge sourcing	-0.024 (0.03)	-0.015 (0.029)	-0.005 (0.029)	-0.043 (0.03)
Public funding	0.028 (0.028)	0.047 (0.026)	0.047 (0.026)	0.030 (0.028)
Incremental innovation	0.000 (0.021)	-0.013 (0.030)	0.001 (0.019)	-0.029 (0.021)

Radical innovation	-0.012 (0.023)	-0.026 (0.026)	-0.016 (0.026)	-0.034 (0.023)
Exports	0.055** (0.024)	0.074*** (0.020)	0.074*** (0.020)	0.055** (0.024)
Lack of employees	0.040* (0.021)	0.035 (0.018)	0.034 (0.018)	0.040* (0.021)
Sales per person	0.033** (0.013)	0.038*** (0.010)	0.038*** (0.010)	0.032** (0.013)
Number of external knowledge sources	0.000 (0.005)	0.001 (0.004)	0.001 (0.004)	0.000 (0.005)
Machine	-0.026 (0.024)	-0.019 (0.021)	-0.026 (0.022)	-0.037 (0.024)
External knowledge	-0.093* (0.051)	-0.151*** (0.047)	-0.151*** (0.047)	-0.094* (0.051)
Trainings	0.084*** (0.029)	0.076*** (0.026)	0.073*** (0.026)	0.083*** (0.029)
Design	-0.011 (0.036)	-0.006 (0.033)	-0.00 (0.033)	-0.012 (0.036)
Market research	0.039 (0.026)	0.032 (0.023)	0.035 (0.023)	0.036 (0.026)
Product innovation		0.024 (0.038)		0.052 (0.43)
Process innovation			0.030 (0.0211)	0.045 (0.25)
Neither				-0.018 (0.63)
Sector and year dummies	Yes	Yes	Yes	Yes
R^2	0.012	0.018	0.012	0.012
LR-chi^2	246.59	533.71	535.36	251.23
p>chi^2	0.000	0.000	0.000	0.000
No of obs.	15331	21334	21334	15331

Notes. Model 1, 2, 3 and 4 stands for the regression models using set of control variables for each outcome, namely external knowledge sourcing/cooperation, product/process innovation and open innovation, respectively. Significance levels are ***p < 1%, **p < 5%, *p < 10%.
Source: authors' calculations from PITEC

c) Treatment group 3

	Model 1	Model 2	Model 3	Model 4
	Coefficient (Std. error)	Coefficient (Std. error)	Coefficient (Std. error)	Coefficient (Std. error)
Group	0.077** (0.030)	0.079*** (0.025)	0.079*** (0.025)	0.078** (0.030)
Share of women	0.132 (0.064)	0.117** (0.050)	0.117** (0.050)	0.131** (0.064)
Employment	0.045*** (0.010)	0.058*** (0.008)	0.058*** (0.008)	0.045*** (0.010)
Employment in R&D	0.007* (0.015)	-0.006* (0.014)	-0.006* (0.014)	0.008 (0.015)
Internal R&D	-0.111** (0.043)	-0.032** (0.035)	-0.032** (0.035)	-0.120** (0.045)
External R&D	0.019 (0.027)	0.012 (0.024)	0.012 (0.024)	0.020 (0.027)
Foreign ownership	-0.056 (0.371)	-0.056 (0.032)	-0.056 (0.032)	-0.056 (0.371)
Cooperation	0.046 (0.028)	0.041 (0.024)	0.040 (0.024)	0.045 (0.028)
External knowledge sourcing	-0.016 (0.042)	-0.019 (0.033)	-0.020 (0.033)	-0.045 (0.059)
Public funding	0.030 (0.032)	0.038 (0.030)	0.038 (0.030)	0.030 (0.032)
Incremental innovation	-0.025 (0.026)	-0.024 (0.034)	-0.023 (0.022)	-0.028 (0.038)

Radical innovation	0.002 (0.026)	0.000 (0.030)	0.002 (0.023)	0.000 (0.034)
Exports	-0.020** (0.028)	0.004 (0.023)	0.004 (0.023)	-0.020 (0.028)
Lack of employees	0.046* (0.025)	0.021 (0.021)	0.021 (0.021)	0.047* (0.025)
Sales per person	0.020 (0.014)	0.023** (0.012)	0.023** (0.012)	0.020 (0.014)
Number of external knowledge sources	-0.007 (0.006)	-0.002 (0.005)	-0.002 (0.005)	-0.00 (0.006)
Machine	-0.068** (0.028)	-0.076*** (0.025)	-0.077*** (0.025)	-0.072** (0.029)
External knowledge	-0.015 (0.058)	-0.054 (0.052)	-0.054 (0.052)	-0.01 (0.058)
Trainings	0.084** (0.034)	0.077** (0.030)	0.077** (0.031)	0.083** (0.034)
Design	-0.013 (0.042)	0.023 (0.038)	0.023 (0.038)	-0.014 (0.042)
Market research	0.045 (0.030)	0.047* (0.027)	0.047* (0.027)	0.046 (0.030)
Product innovation		0.002 (0.043)		0.003 (0.050)
Process innovation			0.005 (0.024)	0.013 (0.029)
Neither				-0.049 (0.073)
Sector and year dummies	Yes	Yes	Yes	Yes
R^2	0.012	0.013	0.013	0.012
LR-chi^2	191.99	287.37	287.41	192.67
p>chi^2	0.012	0.000	0.000	0.000
No of obs.	11712	16716	16716	11712

Notes. Model 1, 2, 3 and 4 stands for the regression models using set of control variables for each outcome, namely external knowledge sourcing/cooperation, product/process innovation and open innovation, respectively. Significance levels are ***p < 1%, **p < 5%, *p < 10%.
Source: authors' calculations from PITEC

Table 14

Probit regression results for increase in Blau index

a) Treatment group 1

	Model 1 Coefficient (Std. error)	Model 2 Coefficient (Std. error)	Model 3 Coefficient (Std. error)	Model 4 Coefficient (Std. error)
Group	0.048** (0.024)	0.040* (0.021)	0.040* (0.021)	0.048** (0.024)
Share of women	0.292*** (0.051)	0.244*** (0.041)	0.243*** (0.041)	0.291*** (0.051)
Employment	0.025*** (0.008)	0.056*** (0.007)	0.055*** (0.007)	0.025*** (0.008)
Employment in R&D	0.014 (0.013)	-0.004 (0.011)	-0.004 (0.011)	0.015 (0.013)
Internal R&D	-0.099** (0.035)	-0.015 (0.028)	-0.014 (0.028)	-0.099*** (0.036)
External R&D	0.013 (0.022)	0.012 (0.020)	0.013 (0.020)	0.013 (0.022)
Foreign ownership	-0.056 (0.030)	-0.051* (0.026)	-0.052* (0.026)	-0.056* (0.030)
Cooperation	0.003 (0.022)	0.006 (0.020)	0.048 (0.020)	0.003 (0.023)

External knowledge sourcing	-0.019 (0.035)	0.007 (0.027)	0.006 (0.027)	-0.023 (0.049)
Public funding	0.029 (0.026)	0.037 (0.024)	0.037 (0.024)	0.030 (0.026)
Incremental innovation	0.001 (0.021)	-0.015 (0.028)	0.006 (0.018)	-0.027 (0.031)
Radical innovation	-0.009 (0.021)	-0.025 (0.024)	-0.010 (0.019)	-0.030 (0.027)
Exports	0.028* (0.023)	0.037* (0.019)	0.037** (0.019)	0.027* (0.023)
Lack of employees	0.050** (0.020)	0.041** (0.017)	0.041** (0.017)	0.051** (0.020)
Sales per person	0.026** (0.012)	0.030*** (0.010)	0.029*** (0.010)	0.026** (0.012)
Number of external knowledge sources	-0.001 (0.005)	-0.001 (0.004)	-0.001 (0.004)	-0.001 (0.005)
Machine	-0.044* (0.023)	-0.037* (0.020)	-0.042** (0.021)	-0.049** (0.024)
External knowledge	-0.079 (0.049)	-0.120*** (0.044)	-0.119*** (0.044)	-0.080 (0.049)
Trainings	0.077*** (0.027)	0.072*** (0.025)	0.070*** (0.025)	0.076*** (0.028)
Design	-0.011 (0.034)	-0.004 (0.031)	-0.005 (0.031)	-0.012 (0.034)
Market research	0.038 (0.024)	0.034 (0.022)	0.038* (0.022)	0.035 (0.025)
Product innovation		0.036 (0.035)		0.049 (0.041)
Process innovation			0.021 (0.020)	0.026 (0.024)
Neither				0.005 (0.060)
Sector and year dummies	Yes	Yes	Yes	Yes
R^2	0.009	0.011	0.011	0.010
LR-chi^2	223.10	379.76	379.85	225.70
p>chi^2	0.000	0.000	0.000	0.000
No of obs.	17246	24303	24303	17246

Notes. Model 1, 2, 3 and 4 stands for the regression models using set of control variables for each outcome, namely external knowledge sourcing/cooperation, product/process innovation and open innovation, respectively. Significance levels are ***p < 1%, **p < 5%, *p < 10%.
Source: authors' calculations from PITEC

b) Treatment group 2

	Model 1 Coefficient (Std. error)	Model 2 Coefficient (Std. error)	Model 3 Coefficient (Std. error)	Model 4 Coefficient (Std. error)
Group	0.046* (0.025)	0.043* (0.021)	0.044** (0.022)	0.046* (0.025)
Share of women	0.362*** (0.054)	0.301*** (0.044)	0.301*** (0.044)	0.362*** (0.054)
Employment	0.046*** (0.009)	0.079*** (0.007)	0.079*** (0.007)	0.046*** (0.009)
Employment in R&D	0.020 (0.013)	0.000 (0.012)	0.000 (0.012)	0.021 (0.013)
Internal R&D	-0.085** (0.037)	-0.003 (0.030)	-0.001 (0.030)	-0.086** (0.038)
External R&D	0.022 (0.023)	0.026 (0.020)	0.026 (0.021)	0.022 (0.023)

Foreign ownership	-0.055* (0.031)	-0.049* (0.027)	-0.050* (0.027)	-0.056* (0.031)
Cooperation	0.008 (0.023)	0.011 (0.021)	0.010 (0.021)	0.005 (0.024)
External knowledge sourcing	-0.029 (0.036)	0.011 (0.029)	0.010 (0.029)	-0.040 (0.051)
Public funding	0.027 (0.027)	0.039 (0.025)	0.040 (0.025)	0.029 (0.027)
Incremental innovation	0.00 (0.022)	-0.006 (0.029)	0.006 (0.019)	-0.022 (0.032)
Radical innovation	-0.014 (0.022)	-0.023 (0.025)	-0.014 (0.020)	-0.033 (0.028)
Exports	0.056** (0.023)	0.073* (0.020)	0.073*** (0.020)	0.056** (0.023)
Lack of employees	0.046** (0.021)	0.042** (0.018)	0.041** (0.018)	0.046** (0.021)
Sales per person	0.033*** (0.012)	0.037*** (0.010)	0.037*** (0.010)	0.033*** (0.012)
Number of external knowledge sources	-0.000 (0.005)	-0.000 (0.004)	-0.000 (0.004)	-0.000 (0.005)
Machine	-0.040* (0.024)	-0.023* (0.021)	-0.028 (0.022)	-0.047* (0.025)
External knowledge	-0.092* (0.051)	-0.136*** (0.046)	-0.136*** (0.046)	-0.093* (0.051)
Trainings	0.084*** (0.028)	0.076*** (0.026)	0.075*** (0.026)	0.083*** (0.028)
Design	-0.014 (0.036)	-0.011 (0.032)	-0.012 (0.032)	-0.015 (0.036)
Market research	0.036 (0.025)	0.026 (0.023)	0.028 (0.023)	0.033 (0.025)
Product innovation		0.021 (0.037)		0.045 (0.042)
Process innovation			0.022 (0.021)	0.029 (0.025)
Neither				-0.006 (0.063)
Sector and year dummies	Yes	Yes	Yes	Yes
R^2	0.013	0.01	0.018	0.013
LR-chi^2	285.05	534.03	534.82	287.52
p>chi^2	0.000	0.000	0.000	0.000
No of obs.	15834	21998	21998	15834

Notes. Model 1, 2, 3 and 4 stands for the regression models using set of control variables for each outcome, namely external knowledge sourcing/cooperation, product/process innovation and open innovation, respectively. Significance levels are ***p < 1%, **p < 5%, *p < 10%.
Source: authors' calculations from PITEC

c) Treatment group 3

	Model 1	Model 2	Model 3	Model 4
	Coefficient (Std. error)	Coefficient (Std. error)	Coefficient (Std. error)	Coefficient (Std. error)
Group	0.070* (0.042)	0.032(0.034)	0.032(0.034)	0.069* (0.042)
Share of women	-0.008 (0.081)	0.022 (0.061)	0.022 (0.061)	-0.010 (0.081)
Employment	-0.069*** (0.014)	-0.035*** (0.010)	-0.037*** (0.010)	-0.069*** (0.014)
Employment in R&D	-0.053** (0.023)	-0.070*** (0.020)	-0.070*** (0.020)	-0.054* (0.023)

Internal R&D	-0.115* (0.058)	-0.016 (0.045)	-0.015 (0.045)	-0.105* (0.061)
External R&D	-0.038 (0.038)	-0.042 (0.033)	-0.043 (0.033)	-0.038 (0.038)
Foreign ownership	-0.080 (0.056)	-0.074 (0.046)	-0.074 (0.046)	-0.080 (0.056)
Cooperation	-0.021 (0.038)	-0.021 (0.032)	-0.020 (0.032)	-0.023 (0.038)
External knowledge sourcing	0.025 (0.057)	0.001 (0.042)	0.003 (0.043)	0.054 (0.083)
Public funding	0.041 (0.045)	0.027 (0.040)	0.026 (0.040)	0.043 (0.045)
Incremental innovation	-0.018 (0.035)	-0.064 (0.045)	0.008 (0.029)	-0.068 (0.052)
Radical innovation	0.005 (0.036)	-0.044 (0.040)	0.003 (0.0305)	-0.035 (0.047)
Exports	-0.117*** (0.038)	-0.109*** (0.031)	-0.108*** (0.031)	-0.117*** (0.038)
Lack of employees	0.075** (0.034)	0.048* (0.027)	0.047* (0.027)	0.076** (0.034)
Sales per person	-0.003 (0.020)	0.026 (0.015)	0.002 (0.015)	-0.003 (0.020)
Number of external knowledge sources	-0.006 (0.008)	-0.003 (0.007)	-0.004 (0.007)	-0.006 (0.008)
Machine	-0.047 (0.040)	-0.093*** (0.033)	-0.097*** (0.034)	-0.048 (0.041)
External knowledge	0.013 (0.086)	-0.038 (0.073)	-0.037 (0.073)	0.013 (0.086)
Trainings	0.034** (0.049)	0.045 (0.042)	0.044 (0.042)	0.035** (0.049)
Design	0.001 (0.059)	0.023 (0.051)	0.020 (0.051)	0.003 (0.060)
Market research	0.048 (0.042)	0.069* (0.037)	0.078** (0.037)	0.041 (0.043)
Product innovation		0.106 (0.058)		0.089 (0.069)
Process innovation			0.015 (0.031)	0.010 (0.039)
Neither				0.057 (0.099)
Sector and year dummies	Yes	Yes	Yes	Yes
R^2	0.035	0.027	0.027	0.035
LR-chi^2	267.72	330.10	326.94	269.74
p>chi^2	0.000	0.000	0.000	0.000
No of obs.	8785	12902	12902	8785

Notes. Model 1, 2, 3 and 4 stands for the regression models using set of control variables for each outcome, namely external knowledge sourcing/cooperation, product/process innovation and open innovation, respectively. Significance levels are ***p < 1%, **p < 5%, *p < 10%.
Source: authors' calculations from PITEC

Proceeding with propensity score matching, dependent variables has been led by t+3 to account for the changes in increased share of women to take place. We get the effect of increase in share of women on likelihood of performing external knowledge sourcing, cooperation and open, product, process innovation activities. **Table 15** presents the result of the matching for each treatment group. In the first treatment group, there is no relationship between increasing share of women and each of 5 outcomes. **Table 15b** demonstrates that firms with an increase between 0 and 10% are 1.1% and 1.4% more likely to engage in product and process innovation, respectively. More than 10% increase in share of women decreases the likelihood to promote open innovation by 1.7% at 10% significance level as seen in **Table 15c**.

When it comes to gender diversity as a treatment variable, positive change in Blau index decreases the likelihood to foster open innovation by 1.1%, as shown in **Table 17a.** Positive association with process innovation is revealed for treatment group of companies with increase in Blau index between 0 and 0.1 at 1% significance level. Having increased gender diversity in the workplace increases the probability of applying process innovation by 1.6%. However, the increased Blau index in this treatment group decreases the likelihood of cooperating by 1.3%. As **Table 17b illustrates**, there is no significant linkage between open innovation and enhanced gender diversity here. **Table 17c** displays that in the treatment group with more than 0.1 increase in Blau index, the probability that firms employ external knowledge sourcing, open, product and process innovation decreases by 2.9%, 4.9%, 4.2% and 1.6%, correspondingly.

Table 16 and 18 demonstrate the findings of propensity score matching for treatment variables with decrease. Negative change in share of women employees and in Blau index decreases the probability to adapt open innovation by 1.1% and 1.3%, respectively.

Table 15

PSM results with increase in share of women as treatment variable

a) Treatment group 1

Variable	Sample	Treated	Controls	Difference (ATT)	Std. Err.	Significance
Product Innovation (t+3)	Unmatched	0.450	0.432	0.017	0.006	***
ATT	Matched	0.4505	0.449	0.001	0.007	NS
Process Innovation (t+3)	Unmatched	0.451	0.421	0.030	0.006	***
ATT	Matched	0.451	0.447	0.004	0.007	NS
External knowledge sourcing (t+3)	Unmatched	0.844	0.846	-0.002	0.006	NS
ATT	Matched	0.844	0.850	-0.006	0.006	NS
Cooperation (t+3)	Unmatched	0.433	0.421	0.012	0.008	NS
ATT	Matched	0.433	0.441	-0.008	0.009	NS
Open Innovation (t+3)	Unmatched	0.609	0.617	-0.076	0.008	NS
ATT	Matched	0.609	0.618	-0.088	0.009	NS

Notes. Propensity score matching results are found based on Table 13a. The number of observations is the same with the regression.
Source: authors' calculations from PITEC

b) Treatment group 2

Variable	Sample	Treated	Controls	Difference (ATT)	Std. Err.	Significance
Product Innovation (t+3)	Unmatched	0.472	0.432	0.398	0.007	***
ATT	Matched	0.472	0.461	0.011	0.008	*
Process Innovation (t+3)	Unmatched	0.472	0.421	0.051	0.007	***
ATT	Matched	0.472	0.457	0.014	0.008	**
External knowledge sourcing (t+3)	Unmatched	0.851	0.846	0.004	0.006	NS
ATT	Matched	0.851	0.852	-0.001	0.007	NS
Cooperation (t+3)	Unmatched	0.444	0.421	0.023	0.008	***
ATT	Matched	0.444	0.446	-0.002	0.009	NS
Open Innovation (t+3)	Unmatched	0.625	0.617	0.099	0.008	NS
ATT	Matched	0.625	0.620	0.006	0.009	NS

Notes. Propensity score matching results are found based on Table 13b. The number of observations is the same with the regression.
Source: authors' calculations from PITEC

c) Treatment group 3

Variable	Sample	Treated	Controls	Difference (ATT)	Std. Err.	Significance
Product Innovation (t+3)	Unmatched	0.440	0.432	0.008	0.009	NS
ATT	Matched	0.440	0.433	0.007	0.009	NS
Process Innovation (t+3)	Unmatched	0.444	0.421	0.023	0.008	***
ATT	Matched	0.444	0.444	-0.00	0.009	NS
External knowledge sourcing (t+3)	Unmatched	0.840	0.846	-0.006	0.007	NS
ATT	Matched	0.840	0.845	-0.004	0.008	NS
Cooperation (t+3)	Unmatched	0.435	0.421	0.014	0.009	NS
ATT	Matched	0.435	0.448	-0.013	0.010	NS
Open Innovation (t+3)	Unmatched	0.593	0.617	-0.024	0.009	***
ATT	Matched	0.593	0.610	-0.017	0.011	*

Notes. Propensity score matching results are found based on Table 13c. The number of observations is the same with the regression.
Source: authors' calculations from PITEC

Table 16

PSM results with decrease in share of women as treatment variable

Variable (Dec)	Sample	Treated	Controls	Difference (ATT)	Std. Err.	Significance
Product Innovation (t+3)	Unmatched	0.460	0.432	0.028	0.006	***
ATT	Matched	0.460	0.454	0.006	0.007	NS
Process Innovation (t+3)	Unmatched	0.452	0.421	0.031	0.006	***
ATT	Matched	0.452	0.450	0.002	0.007	NS
External knowledge sourcing (t+3)	Unmatched	0.851	0.846	0.004	0.005	NS
ATT	Matched	0.851	0.858	-0.007	0.006	NS
Cooperation (t+3)	Unmatched	0.426	0.421	0.006	0.007	NS
ATT	Matched	0.426	0.430	-0.003	0.008	NS
Open Innovation (t+3)	Unmatched	0.615	0.617	-0.002	0.007	NS
ATT	Matched	0.615	0.626	-0.011	0.008	*

Notes. The number of observations is 26,617 for product and process innovation, 18,990 for external knowledge sourcing and open innovation.
Source: authors' calculations from PITEC

Table 17

PSM results with Blau index as treatment variable

a) Treatment group 1

Variable	Sample	Treated	Controls	Difference (ATT)	Std. Err.	Significance
Product Innovation (t+3)	Unmatched	0.451	0.432	0.019	0.006	***
ATT	Matched	0.451	0.447	0.004	0.007	NS
Process Innovation (t+3)	Unmatched	0.451	0.420	0.031	0.006	***
ATT	Matched	0.451	0.443	0.008	0.007	NS
External knowledge sourcing (t+3)	Unmatched	0.844	0.846	-0.002	0.006	NS
ATT	Matched	0.844	0.850	-0.006	0.006	NS
Cooperation (t+3)	Unmatched	0.429	0.421	0.009	0.008	NS
ATT	Matched	0.429	0.433	-0.005	0.009	NS
Open Innovation (t+3)	Unmatched	0.610	0.617	-0.007	0.007	NS
ATT	Matched	0.621	0.621	-0.011	0.008	*

Notes. Propensity score matching results are found based on Table 14a. The number of observations is the same with the regression.
Source: authors' calculations from PITEC

b) Treatment group 2

Variable	Sample	Treated	Controls	Difference (ATT)	Std. Err.	Significance
Product Innovation (t+3)	Unmatched	0.473	0.432	0.041	0.007	***
ATT	Matched	0.473	0.464	0.008	0.008	NS
Process Innovation (t+3)	Unmatched	0.470	0.420	0.050	0.007	***
ATT	Matched	0.470	0.453	0.016	0.008	***
External knowledge sourcing (t+3)	Unmatched	0.851	0.846	0.005	0.006	NS
ATT	Matched	0.851	0.858	-0.006	0.006	NS
Cooperation (t+3)	Unmatched	0.437	0.420	0.016	0.008	**
ATT	Matched	0.437	0.450	-0.013	0.009	*
Open Innovation (t+3)	Unmatched	0.626	0.617	0.009	0.008	NS
ATT	Matched	0.626	0.624	0.002	0.009	NS

Notes. Propensity score matching results are found based on Table 14b. The number of observations is the same with the regression.
Source: authors' calculations from PITEC

c) Treatment group 3

Variable	Sample	Treated	Controls	Difference (ATT)	Std. Err.	Significance
Product Innovation (t+3)	Unmatched	0.346	0.432	-0.086	0.011	***
ATT	Matched	0.346	0.387	-0.042	0.013	***
Process Innovation (t+3)	Unmatched	0.359	0.420	-0.061	0.011	***
ATT	Matched	0.359	0.375	-0.016	0.013	*
External knowledge sourcing (t+3)	Unmatched	0.804	0.846	-0.042	0.011	***
ATT	Matched	0.804	0.833	-0.029	0.013	***
Cooperation (t+3)	Unmatched	0.381	0.420	-0.039	0.014	***
ATT	Matched	0.381	0.473	0.008	0.016	NS
Open Innovation (t+3)	Unmatched	0.514	0.617	-0.102	0.014	***
ATT	Matched	0.514	0.563	-0.049	0.016	***

Notes. Propensity score matching results are found based on Table 14c. The number of observations is the same with the regression.
Source: authors' calculations from PITEC

Table 18

PSM results with decrease in Blau index

Variable	Sample	Treated	Controls	Difference (ATT)	Std. Err.	Significance
Product Innovation (t+3)	Unmatched	0.446	0.432	0.014	0.007	***
ATT	Matched	0.446	0.444	0.002	0.007	NS
Process Innovation (t+3)	Unmatched	0.438	0.420	0.018	0.007	***
ATT	Matched	0.438	0.421	0.011	0.007	*
External knowledge sourcing (t+3)	Unmatched	0.849	0.846	0.003	0.006	NS
ATT	Matched	0.849	0.854	-0.006	0.006	NS
Cooperation (t+3)	Unmatched	0.419	0.420	-0.001	0.007	NS
ATT	Matched	0.419	0.431	-0.012	0.009	*
Open Innovation (t+3)	Unmatched	0.609	0.617	-0.008	0.008	NS
ATT	Matched	0.609	0.622	-0.013	0.009	*

Notes. The number of observations is 16,350 for open innovation, cooperation and external sourcing, 23,229 for product and process innovation.
Source: authors' calculations from PITEC

Moreover, we extend our analysis to explore this effect for manufacturing and service industry firms shown in **Table 19.** In manufacturing firms, the treatment group with more than 0.1 increase in gender diversity has a significant negative association with open innovation. Additionally, T**able 19a** demonstrates that the probability to apply open innovation decrease for manufacturing firms that have more than 0% or more than 10% increase in their proportion of women employees. However, for more than 10% increase in share women, utilization of open innovation in service industry firms declines more compared to manufacturing companies. Service industry firms that increase the share of women more than 10% have 3.1% less likelihood to perform open innovation as seen in **Table 19c**. For gender diversity as treatment variable in service industry firms, only significant result is shown for 3rd treatment group. Firms with more than 0.1 increase in Blau index the probability of leveraging open innovation activities decreases by 4.5%.

Table 20 shows us the balance between increased share of women and covariates before and after matching. We can see that the difference between the means of treated and control group has been decreased after matching indicating good balancing results.

Table 20

Average balancing property tests after propensity score matching

		Mean		Bias	t-test	
	Unmatched Matched	Treated	Matched Control	%	t-value	p-value
Group	U	0.479	0.424	11.1	5.83	0.000
	M	0.479	0.481	-0.3	-0.16	0.870
Share of women	U	0.304	0.287	7.4	3.87	0.000
	M	0.304	0.308	-1.8	-0.83	0.408
Employment	U	4.543	4.302	14	7.33	0.000
	M	4.543	4.569	-1.6	-0.72	0.470
Employment in R&D	U	1.414	1.389	1.8	0.96	0.338
	M	1.414	1.407	0.5	0.22	0.829
Internal R&D	U	0.682	0.717	-7.7	-4.06	0.000
	M	0.682	0.679	0.8	0.35	0.726
External R&D	U	0.385	0.371	2.8	1.49	0.135
	M	0.385	0.390	-1.0	-0.45	0.652
Foreign ownership	U	0.162	0.154	2.1	1.12	0.264
	M	0.162	0.165	-0.8	-0.36	0.723
Cooperation	U	0.419	0.396	4.7	2.48	0.013
	M	0.419	0.425	-1.2	-0.58	0.561

External knowledge sourcing	U	0.843	0.857	-4.0	-2.13	0.033
	M	0.843	0.840	0.6	0.29	0.771
Public funding	U	0.278	0.273	1.0	0.55	0.583
	M	0.278	0.278	0.0	1.01	0.994
Incremental innovation	U	0.518	0.534	-3.2	-1.69	0.091
	M	0.518	0.519	-0.2	-0.09	0.926
Radical innovation	U	0.422	0.425	-0.7	-0.37	0.713
	M	0.422	0.412	1.9	0.90	0.368
Exports	U	0.498	0.516	-3.5	-1.83	0.067
	M	0.498	0.497	0.2	0.11	0.915
Lack of employees	U	0.415	0.408	1.5	0.80	0.421
	M	0.415	0.417	-0.4	-0.19	0.846
Sales per person	U	11.9	11.873	2.6	1.37	0.170
	M	11.9	11.89	0.9	0.44	0.660
Number of external sources	U	3.640	3.697	-2.1	-1.09	0.276
	M	3.640	3.577	2.3	1.08	0.281
Machine	U	0.274	0.276	-0.3	-0.18	0.855
	M	0.274	0.274	-0.2	-0.08	0.937
External knowledge	U	0.051	0.045	3.2	1.68	0.094
	M	0.051	0.051	0.0	0.00	1.000
Trainings	U	0.201	0.173	7.3	3.87	0.000
	M	0.201	0.198	1.0	0.45	0.651
Design	U	0.104	0.098	1.9	1.02	0.309
	M	0.104	0.105	-0.5	-0.21	0.835
Market research	U	0.279	0.266	2.8	1.48	0.138
	M	0.279	0.281	-0.6	-0.29	0.770
Product Innovation	U	0.683	0.698	-3.3	-1.75	0.081
	M	0.683	0.684	-0.4	-0.18	0.855
Process Innovation	U	0.707	0.706	0.2	0.10	0.924
	M	0.707	0.713	-1.2	-0.56	0.574
Neither	U	0.104	0.089	5.1	2.72	0.007
	M	0.104	0.103	0.1	0.06	0.954

Notes. Number of observations is same with regression Model 4 on Table 13c.
Source: authors' calculations from PITEC

3.2 Sensitivity Analysis

We perform the following robustness tests: changing number of neighbors, threshold and matching method.

First, we adjust the number of nearest neighbors to match. The results of matching suggest consistency for pairs ranging from one to seven for first and second treatment groups in case of increased share of women as treatment variable. For increase more than 10%, starting from 2 neighbors the significant decreasing relationship continues. With regard to the increase in Blau index more than 0, the significant negative association with open innovation persists only for 2 and 3 neighbors. Similarly, we still obtain significant negative relationship for more than 0.1 positive change in gender diversity after changing the number of neighbors.

Then changing the threshold for the increase of women employees and gender diversity indicator, we find that significance of previous results on the open innovation outcome still holds. Estimating propensity scores with 5%, 15% threshold shows that both treatment groups with more than 5% and 15% increase have significant reverse impact on open innovation. We also altered the threshold for Blau index to 0.05 and 0.2 levels. Increasing gender diversity between 0 and 0.05 or 0 and 0.2 doesn't change the probability of engaging in open activities. An increase over those threshold levels significantly decreases the likelihood of incorporating open innovation practices to workplace.

Furthermore, we repeat the analysis using kernel matching instead of nearest neighbor. The findings on open innovation follow the same pattern. Firms with more than 10% increased share of women and more than 0.1 increase in gender diversity are 1.9% and 5.6% less likely to participate in open innovation, accordingly, as **Table 21** demonstrates.
The results show that there are no significant changes in the main outcome of our analysis after conducting sensitivity tests. Therefore, the findings presented in the previous section are considered to be robust.

3.3 Discussion

While theoretical expectations lean towards a positive effect of gender diversity on innovation, empirical findings do not demonstrate a general consensus on this topic. Given that feminine behaviour aligns with the main characteristics of open innovation (Wikhamn & Knights, 2013), we sought to explore the possible positive effect of a diverse environment. Nonetheless, our study revealed a more complex non-linear relationship between gender diversity and innovation performance. For companies with an increase in a share of women up to 10% and the Blau index up to 0.1, difference of means test results suggests a positive association with product, process, and open innovation. However, increasing the share of women above a 10% threshold is negatively associated with open innovation performance. On the other hand, decreasing the share of women in the organizations also has a negative relationship with open innovation. Thus, excluding the women representation can be adversely linked with open innovation.

The observed patterns of non-linearity could indicate several factors, underscoring the importance of performing a cost and benefit analysis. At first, increased gender diversity may present the broader information pool and enhanced problem solving, improving the innovative thinking. Subsequent negative relationship above a 10% threshold might lead to challenges in team communication and coherence, causing disruption in firm dynamics and processes. Potential positive effect might take more time to materialize for large scale changes in team composition. Alternatively, certain tasks benefit from a more homogenous workforce. In this regard, studies by Masso et al. (2022), and Masso and Vahter (2020) demonstrate that innovation is associated with high performing firms, where men tend to self-select themselves, resulting in a greater demand for men.

One implication is that there might be an optimal level of gender diversity that positively affects innovation outcomes. Thus, potential drawbacks of increased female representation can outweigh the benefits it brings to innovation and can cause disruption in activities after some point. Hence, management should strive for a level of gender diversity that fosters inclusion keeping in mind the tradeoff between the costs and positive externalities.

A novelty of our research is extending the analysis by considering the decrease in the share of women and decrease in the Blau index. The negative association with both specifications confirms the idea that balanced teams are crucial for adopting innovation strategies. Limiting the proportion of women can result in a lack of diverse perspectives, hindering the creation of innovative solutions. Additionally, given that Spain is a relatively conservative country, one could argue that the institutional context has a role to play in this matter too, as discussed by Adams et al. (2023).

These insights suggest the importance of gender inclusivity for policy makers. Focusing on balanced teams, rather than maximizing the share of women, creating the environment suitable for active participation of all members, reassessing the corporate culture to integrate women are some examples. Investing in these policies can help maintain the benefits women can bring to innovation in the long run. Consequently, the optimal threshold for gender diversity can change based on societal and institutional dynamics.

Conclusion

The aim of our paper was to investigate the relationship between the gender diversity of workforce and their innovation performance with a focus on product, process and open innovations. Using a panel dataset of Spanish firms in 2004-2016, we utilized propensity score matching method to analyze the relationship. Since open innovation is defined as combination of internal R&D and external knowledge sourcing or collaboration, we investigate them separately and together.

We find that, on average, firms that increased the share of women employees in yearly basis displays different outcomes than the ones with no increase. Companies with increase up to 10% have more likelihood to use product and process innovation. More than 10% increase in the share of women employees associates negatively with open innovation. The probability to practice open innovation decreases by 1.7% at 10% significance level. However, there is no relationship between the increase in share of women and collaboration, external knowledge sources, product and process innovation in this treatment group. Industry level results indicate that for manufacturing and service sectors the likelihood to openly innovate negatively associate with between 0 and 10% increase in share of women.

When it comes to gender diversity in firms as a treatment variable that was assigned based on a change Blau index, there is a negative linkage with open innovation if we take companies with a positive change in our treatment group. Gender diverse firms with more than 0.1 change in Blau are less inclined to engage in external knowledge and product innovation activities. Additionally, the probability of those companies using open innovation decreases by 4.9%. In manufacturing firms, the negative relationship is significant for open innovation in case of the increase more than 0.1 in Blau. Similarly, more than 0.1 increase in gender diversity in service industry firms shows negative significant change in the probability of employing open innovation.

Our paper has several limitations. Firstly, the scope of our analysis is limited to companies in Spain from 2004 to 2016. Our results are valid for surviving firms engaging in innovative activities. Secondly, in addition to the selected variables, there might be unobserved factors that can affect outcome. Subsequent research could explore this topic in greater depth by considering alternative combinations of open innovation beyond the R&D and external linkages. Exploring the relationship with longer time span (t+6) for the changes in outcome variables to take place after an increase in share of women presents another avenue for future research. As our exploration of gender diversity focused on the overall workforce level, it is worth considering that the effects may differ across various positions and departments, such as R&D and corporate boards. Spain is an interesting example to examine the effect of corporate gender quotas.

This papers brings interesting contributions to both practical and theoretical aspects. Our results suggest that the increase in female representation should be balanced. Researching both increase and decrease of gender diversity brings novelty to the literature and is valuable for policy implications.

References

Adams, K., Attah-Boakye, R., Yu, H., Johansson, J., & Njoya, E. T. (2023). Female board representation and coupled open innovation: Evidence from emerging market multinational enterprises. *Technovation, 124,* 102749. https://doi.org/10.1016/j.technovation.2023.102749

Almor, T., Bazel-Shoham, O., & Lee, S. M. (2022). The dual effect of board gender diversity on R&D investments. *Long Range Planning*, *55*(2), 101884. https://doi.org/10.1016/j.lrp.2019.05.004

Alsos, G. A., Ljunggren, E., & Hytti, U. (2013). Gender and innovation: State of the art and a research agenda. *International Journal of Gender and Entrepreneurship*, *5*(3), 236–256. https://doi.org/10.1108/IJGE-06-2013-0049

Becker, B., & Vahter, P. (2022). Gender Diversity in R&D Teams and Firms' Eco-innovation Strategies. *Academy of Management Proceedings, 2022*(1), 16905. https://doi.org/10.5465/AMBPP.2022.16905abstract

Bertrand, M. (2011). New Perspectives on Gender. In Handbook of Labor Economics. *Elsevier, 4,*1543–1590 . https://doi.org/10.1016/S0169-7218(11)02415-4

Bianchini, S., Pellegrino, G., & Tamagni, F. (2018). Innovation complementarities and firm growth. *Industrial and Corporate Change, 27*(4), 657–676. https://doi.org/10.1093/icc/dty008

Caliendo, M., & Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys, 22(*1), 31–72. https://doi.org/10.1111/j.1467-6419.2007.00527.x

Cassiman, B., & Veugelers, R. (2006). In Search of Complementarity in Innovation Strategy: Internal R&D and External Knowledge Acquisition. *Management Science.* https://doi.org/10.1287/mnsc.1050.0470

Chen, J., Leung, W. S., & Evans, K. P. (2018). Female board representation, corporate innovation and firm performance. *Journal of Empirical Finance, 48,* 236–254. https://doi.org/10.1016/j.jempfin.2018.07.003

Chesbrough, H. (2003). *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Harvard Business School Press, Boston, MA.

Chesbrough, H., Vanhaverbeke, W., & West, J. (2006). *Open Innovation: Researching a New Paradigm*. Oxford University Press, Oxford.

Comi, S., Grasseni, M., Origo, F., & Pagani, L. (2020). Where Women Make a Difference: Gender Quotas and Firms' Performance in Three European Countries. *ILR Review, 73*(3), 768–793. https://doi.org/10.1177/0019793919846450

Croson, R., & Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature, 47(*2), 448–474. https://doi.org/10.1257/jel.47.2.448

Dahlander, L., & Gann, D. M. (2010). How open is innovation? *Research Policy, 39*(6), 699–709. https://doi.org/10.1016/j.respol.2010.01.013

Dezsö, C. L., & Ross, D. G. (2012). Does female representation in top management improve firm performance? A panel data investigation. *Strategic Management Journal*, *33*(9), 1072–1089. https://doi.org/10.1002/smj.1955

Díaz-García, C., González-Moreno, A., & Jose Sáez-Martínez, F. (2013). Gender diversity within R&D teams: Its impact on radicalness of innovation. *Innovation, 15(*2), 149–160. https://doi.org/10.5172/impp.2013.15.2.149

Duch-Brown, N., De Panizza, A., & Rohman, I. K. (2016). *Innovation and productivity in an S&T-intensive sector—Publications Office of the EU.* https://op.europa.eu/en/publication-detail/-/publication/5498f473-4729-11e6-9c64-01aa75ed71a1/language-en

Durst, S., & Ståhle, P. (2013). Success Factors of Open Innovation - A Literature Review. *International Journal of Business Research and Management, 4*(4), 111-131. http://www.cscjournals.org/manuscript/Journals/IJBRM/Volume4/Issue4/IJBRM-154.pdf

Gong, Y., & Girma, S. (2021). Top management gender diversity and performance: In search of threshold effects. *Cambridge Journal of Economics, 45(*1), 109–127. https://doi.org/10.1093/cje/beaa045

Griffin, D., Li, K., & Xu, T. (2021). Board Gender Diversity and Corporate Innovation: *International Evidence. Journal of Financial and Quantitative Analysis, 56*(1), 123–154. https://doi.org/10.1017/S002210901900098X

Joecks, J., Pull, K., & Vetter, K. (2013). Gender Diversity in the Boardroom and Firm Performance: What Exactly Constitutes a "Critical Mass?" *Journal of Business Ethics, 118(*1), 61–72.

Laursen, K., & Salter, A. (2006). Open for innovation: The role of openness in explaining innovation performance among UK manufacturing firms. *Strategic Management Journal*, *27*(2), 131–150. https://doi.org/10.1002/smj.507

Lee, J., & Chung, J. (2022). Women in top management teams and their impact on innovation. *Technological Forecasting and Social Change, 183,* 121883. https://doi.org/10.1016/j.techfore.2022.121883

Love, J. H., Roper, S., & Vahter, P. (2014). Dynamic complementarities in innovation strategies. *Research Policy, 43(*10), 1774–1784. https://doi.org/10.1016/j.respol.2014.05.005

Maida, A., & Weber, A. (2019). Female Leadership and Gender Gap within Firms: Evidence from an Italian Board Reform. https://docs.iza.org/dp12099.pdf

Masso, J., Meriküll, J., & Vahter, P. (2022). The role of firms in the gender wage gap. *Journal of Comparative Economics, 50*(2), 454–473. https://doi.org/10.1016/j.jce.2021.10.001

Masso, J., & Vahter, P. (2020). Innovation as a Firm-Level Factor of the Gender Wage Gap. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.3756783

OECD & Eurostat. (2018). Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation, 4th Edition*. OECD.* https://doi.org/10.1787/9789264304604-en

OECD & Statistical Office of the European Communities. (2005). Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data, 3rd Edition. *OECD.* https://doi.org/10.1787/9789264013100-en

Østergaard, C. R., Timmermans, B., & Kristinsson, K. (2011). Does a different view create something new? The effect of employee diversity on innovation. *Research Policy*, *40*(3), 500–509. https://doi.org/10.1016/j.respol.2010.11.004

Post, C., & Byron, K. (2015). Women on Boards and Firm Financial Performance: A Meta-Analysis. *The Academy of Management Journal, 58(*5), 1546–1571.

Reguera-Alvarado, N., De Fuentes, P., & Laffarga, J. (2017). Does Board Gender Diversity Influence Financial Performance? Evidence from Spain. *Journal of Business Ethics, 141*(2), 337–350. https://doi.org/10.1007/s10551-015-2735-9

Ritter-Hayashi, D., Vermeulen, P., & Knoben, J. (2019). Is this a man's world? The effect of gender diversity and gender equality on firm innovativeness. *PLOS ONE*, *14*(9), e0222443. https://doi.org/10.1371/journal.pone.0222443

Rodriguez, M., Doloreux, D., & Shearmur, R. (2017). Variety in external knowledge sourcing and innovation novelty: Evidence from the KIBS sector in Spain. *Technovation, 68*, 35–43. https://doi.org/10.1016/j.technovation.2017.06.003

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70(*1), 41–55. https://doi.org/10.1093/biomet/70.1.41

Sainani, K. L. (2012). Propensity Scores: Uses and Limitations. *PM&R, 4*(9), 693–697. https://doi.org/10.1016/j.pmrj.2012.07.002

Sastre, J. F. (2015). The impact of R&D teams' gender diversity on innovation outputs. *International Journal of Entrepreneurship and Small Business, 24(*1), 142–162. https://doi.org/10.1504/IJESB.2015.066154

Schmiedeberg, C. (2008). Complementarities of innovation activities: An empirical analysis of the German manufacturing sector. *Research Policy, 37(*9), 1492–1503. https://doi.org/10.1016/j.respol.2008.07.008

Segarra, A., & Teruel, M. (2014). High-growth firms and innovation: An empirical analysis for Spanish firms. *Small Business Economics, 43*(4), 805–821. https://doi.org/10.1007/s11187-014-9563-7

Tahir, S. H., Ullah, M. R., Ahmad, G., Syed, N., & Qadir, A. (2021). Women in Top Management: Performance of Firms and Open Innovation. *Journal of Open Innovation: Technology, Market, and Complexity, 7(*1), 87. https://doi.org/10.3390/joitmc7010087

Teruel, M., & Segarra-Blasco, A. (2017). Gender Diversity, R&D Teams and Patents: An Application to Spanish Firms. *SSRN*, *3076169*. https://doi.org/10.2139/ssrn.3076169

Tonoyan, V., & Boudreaux, C. J. (2023). Gender diversity in firm ownership: Direct and indirect effects on firm-level innovation across 29 emerging economies. *Research Policy, 52*(4), 104716. https://doi.org/10.1016/j.respol.2022.104716

Tsui, A. S., Egan, T. D., & Iii, C. A. O. (1992). Being Different: Relational Demography and Organizational Attachment. *Administrative Science Quarterly, 37*(4), 549. https://doi.org/10.2307/2393472

West, J., Salter, A., Vanhaverbeke, W., & Chesbrough, H. (2014). Open innovation: The next decade. *Research Policy*, *43*(5), 805–811. https://doi.org/10.1016/j.respol.2014.03.001

Wikhamn, R. B., & Knights, D. (2013). Open innovation, gender and the infiltration of masculine discourses. *International Journal of Gender and Entrepreneurship*, *5*(3), 275–297. https://doi.org/10.1108/IJGE-Sep-2012-0041

World Economic Forum (2007). The Global Gender Gap Report 2007. Geneva, Switzerland. https://www3.weforum.org/docs/WEF_GenderGap_Report_2007.pdf

World Economic Forum (2022). The Global Gender Gap Report 2022. Geneva, Switzerland. https://www3.weforum.org/docs/WEF_GGGR_2022.pdf

Xie, L., Zhou, J., Zong, Q., & Lu, Q. (2020). Gender diversity in R&D teams and innovation efficiency: Role of the innovation context. *Research Policy*, *49*(1), 103885. https://doi.org/10.1016/j.respol.2019.103885

Yang, P., Riepe, J., Moser, K., Pull, K., & Terjesen, S. (2019). Women directors, firm performance, and firm risk: A causal perspective. *The Leadership Quarterly, 30(*5), 101297. https://doi.org/10.1016/j.leaqua.2019.05.004

Yarram, S. R., & Adapa, S. (2022). Women on boards, CSR and risk-taking: An investigation of the interaction effects of gender diversity and CSR on business risk. *Journal of Cleaner Production, 378,* 134493. https://doi.org/10.1016/j.jclepro.2022.134493

Zhang, L. (2020). An Institutional Approach to Gender Diversity and Firm Performance. *Organization Science, 31*(2), 439–457. https://doi.org/10.1287/orsc.2019.1297

Appendices

Table 1

Summary of literature review

Article	Data	Method	Result
Xie et al., 2020	26,708 observations from Chinese firms in 2010-2013	OLS and random effect model	Positive relationship between gender diversity and innovation efficiency
Ostergaard et al., 2011	1648 Danish companies in 2003-2005	Logistic regression	Positive impact of the gender diversity on the likelihood of innovation
Ritter-Hayashi et al., 2019	Cross-country analysis of 18,547 firms in 15 developing countries in 2013-2014	Binary logistic regression	Positive effect of having female representation in top management on innovation
Dezsö & Ross, 2012	S&P 1,500 US firms in 1992-2006	Fixed effects regression of Tobin's q	Positive effect of gender diversity in the top management on the firm performance is moderated by innovation
Laursen & Salter, 2006	2707 firms in the UK in 2001	Double-truncated Tobit model	Firms open to external search are more innovative, possibility of curvilinear relationship
Becker & Vahter, 2022	12,000 Spanish firms in 2009-2016	Propensity score matching technique, probit model, fixed effects	Positive impact of gender diversity on the likelihood of using eco-innovation strategy
Chen et al., 2018	1,224 firms from ExecuComp dataset in 1998 - 2006.	OLS regression, PSM, logit model, IVM	Increased share of female directors leads to greater innovation performance
Lee & Chung, 2022	1983 US firms in 2005-2019	2SLS with IVM	Increased number of women in top management increases the quantity of less

			impactful innovations
Teruel & Segarra-Blasco, 2017	4,085 Spanish firms in 2004-2014	Two-step procedure, Heckman, Mills raito, Blau index, 2SLS	Gender diversity in R&D teams has dual effect on patent capacity
Díaz-García et al., 2013	4277 firms from PITEC Spanish data in 2007	Logistic binary regression	Gender diversity in R&D teams is positively linked to radical innovation
Griffin et al., 2021	45 countries, 12,244 firms in 2001-2014	Hierarchial linear model, maximum likelihood	Greater female presence in boards has positive association with the number of patents and innovation efficiency
Tonoyan & Boudreaux, 2023	29 emerging countries in BEEPS dataset during 2012-2016	Hierarchial linear modeling, multilevel regression model	Increased gender diversity in firm ownership increases the probability to innovate
Love et al., 2014	Irish manufacturing firms in 1991-2008	Tobit model, multinomial logit model	Authors have found evidence that the combined use of external and internal knowledge sources is linked to increased innovation performance due to dynamic complementarities.
Adams et al., 2023	183 organisations in 16 emerging countries in 2009-2019	GMM and quantile regressions	Significant positive relationship between the increased percentage of women on corporate boards and open innovation
Joecks et al., 2013	151 listed German companies during 2000-2005	OLS, Random effects regression	U-shaped correlation between gender diversity on boardroom and firm performance: reaching critical mass 30% of female share is beneficial

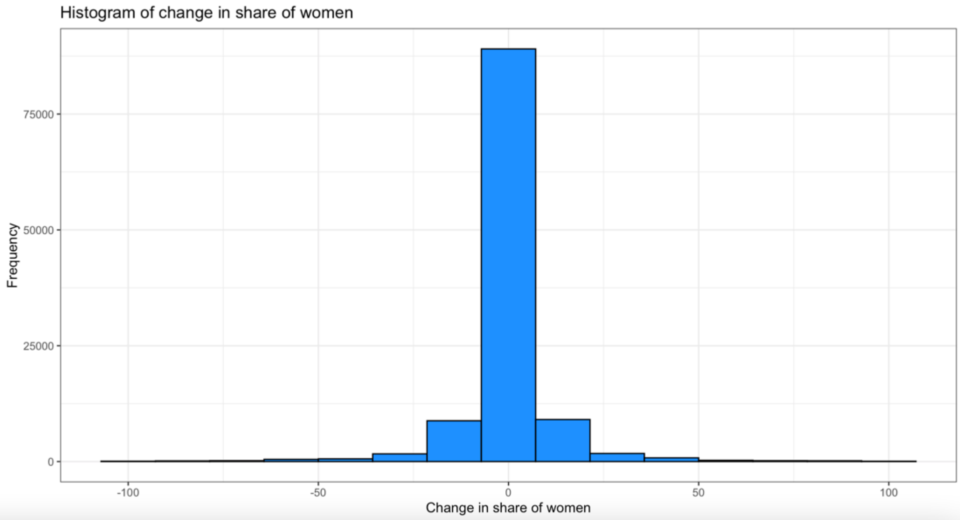

Figure 1. Distribution of change in share of women

Source: own calculations

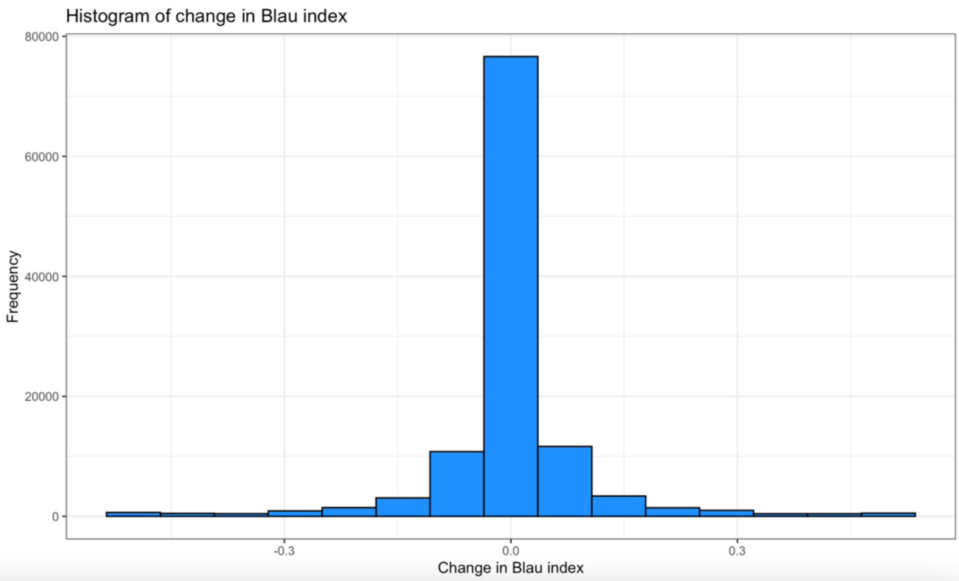

Figure 2. Distribution of change in Blau index
Source: own calculations

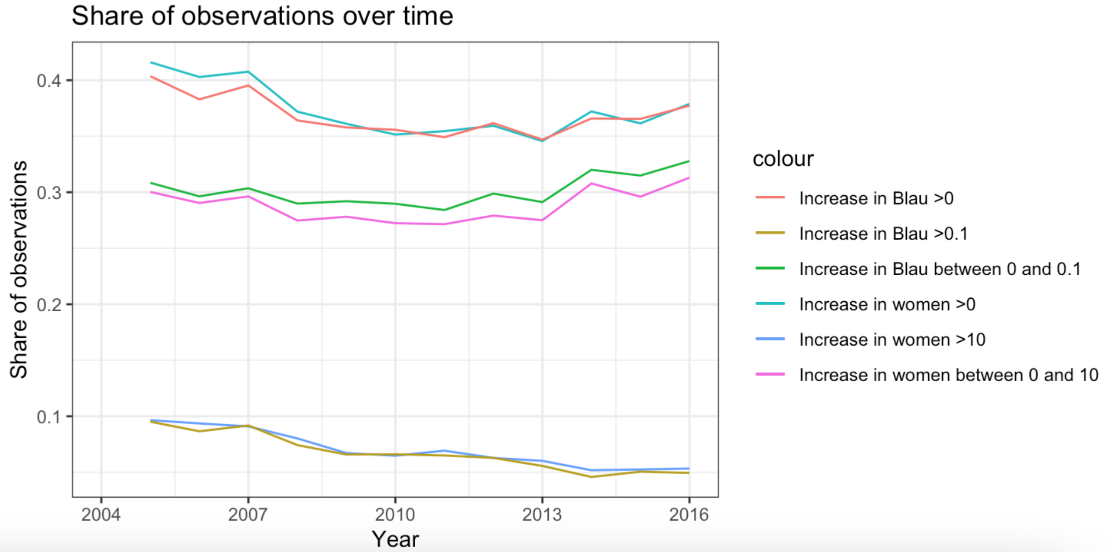

Figure 4. Share of observations in treatment groups.
Source: author's calculations from PITEC

Table 3
Descriptive statistics for industries

	Manufacturing industry			Service industry		
Variables	Obs.	Mean	Std. dev	Obs.	Mean	Std. dev
Number of total employees (ln)	64880	4.026	1.399	53733	4.243	2.042
Number of total employees in R&D (ln)	64880	1.043	1.234	53733	0.783	1.270
Share of women employees	64880	0.262	0.213	53733	0.428	0.265
Share of women employees in R&D	64880	0.145	0.245	53733	0.114	0.221
Increase in share of women	38548	0.542	0.498	30780	0.594	0.491
Increase in share of women in R&D	48940	0.166	0.372	41220	0.139	0.346
Blau index	64880	0.296	0.154	53733	0.350	0.156
Blau index in R&D	64880	0.128	0.188	53733	0.104	0.182
Increase in Blau index	38608	0.539	0.498	30554	0.582	0.493
Group membership	64880	0.400	0.490	53733	0.415	0.493
Internal R&D	64880	0.574	0.494	53733	0.383	0.486
External R&D	64880	0.273	0.445	53733	0.173	0.378
External knowledge sourcing	52941	0.834	0.372	35004	0.787	0.410
Cooperation	52941	0.347	0.476	35005	0.394	0.489
Incremental innovation	64880	0.443	0.497	53733	0.274	0.446
Radical innovation	64880	0.318	0.466	53733	0.226	0.418
Product innovation	64880	0.563	0.496	53733	0.377	0.485
Process innovation	64880	0.549	0.498	53733	0.392	0.488

Open innovation	52941	0.641	0.480	35004	0.531	0.499
Neither	52941	0.103	0.305	35004	0.156	0.363
Foreign ownership	64880	0.159	0.366	53733	0.132	0.339
Public funding	57363	0.144	0.351	49778	0.152	0.359
Exports	64880	0.616	0.486	53733	0.219	0.414
Lack of employees	64880	0.411	0.492	53733	0.324	0.468
Sales per person (ln)	64880	11.965	0.879	53733	11.312	1.345
Number of external knowledge sources	64880	2.862	2.828	53733	2.197	2.855
Machine	64880	0.197	0.398	53733	0.145	0.352
External knowledge acquisition	64880	0.028	0.166	53733	0.030	0.170
Trainings	64880	0.136	0.343	53733	0.130	0.336
Market research	64880	0.215	0.411	53733	0.137	0.344
Design	64880	0.090	0.286	53733	0.055	0.228

Notes. NA values are excluded.
Source: authors' calculations from PITEC

Table 5

Mean t-test. The comparison between companies that have experienced an increase in the proportion of women in their total workforce and those that have not.

	Increase more than 0%			
	Group 0 Mean (SD)	Group 1 Mean (SD)	Difference	Significance
Group membership	0.38 (0.49)	0.44 (0.50)	-0.06	***
Internal R&D	0.46 90.50)	0.48 (0.50)	-0.01	***
External R&D	0.21 (0.40)	0.23 (0.42)	-0.02	***
External knowledge sourcing	0.83 (0.37)	0.85 (0.36)	-0.01	***
Cooperation	0.38 (0.48)	0.40 (0.49)	-0.02	***
Incremental innovation	0.34 (0.47)	0.36 (0.48)	-0.03	***
Radical innovation	0.26 (0.44)	0.28 (0.45)	-0.02	***
Product innovation	0.44 (0.50)	0.48 (0.50)	-0.03	***
Process innovation	0.45 (0.50)	0.50 (0.50)	-0.05	***
Open innovation	0.61 (0.49)	0.61 (0.49)	0.00	NS
Neither	0.11 (0.31)	0.11 (0.31)	0.00	NS
Foreign ownership	0.14 (0.35)	0.15 (0.36)	-0.002	***
Public funding	0.14 (0.34)	0.16 (0.36)	-0.01	***
Exports	0.40 (0.49)	0.42 (0.49)	-0.03	***
Lack of employees	0.37 (0.48)	0.37 (0.48)	0.00	NS
Sales per person	11.68 (1.20)	11.73(1.14)	-0.05	***
Number of external knowledge sources	2.47 (2.86)	2.65 (2.90)	-0.18	***
Machine	0.15 (0.35)	0.17 (0.37)	-0.02	***

External knowledge acquisition	0.02 (0.14)	0.02 (0.15)	0.00	***
Trainings	0.10 (0.30)	0.12 (0.32)	-0.01	***
Market research	0.16 (0.37)	0.17 (0.38)	-0.01	***
Design	0.06 (0.23)	0.06 (0.24)	0.00	**
Number of observations	32,826	42,472		

Notes. Welch Two Sample t-test; NA values are excluded. Standard errors are presented in parentheses.
Source: authors' calculations from PITEC

Table 8
Mean t-test. The comparison between companies with no change and a positive change in their Blau index

	Increase more than 0			
	Group 0 Mean (SD)	Group 1 Mean (SD)	Difference	Significance
Group membership	0.38 (0.49)	0.45 (0.50)	-0.06	***
Internal R&D	0.46 (0.50)	0.48 (0.50)	-0.02	***
External R&D	0.21 (0.40)	0.23 (0.42)	-0.03	***
External knowledge sourcing	0.83 (0.37)	0.85 (0.36)	-0.01	***
Cooperation	0.38 (0.48)	0.40 (0.49)	-0.02	***
Incremental innovation	0.34 (0.47)	0.37 (0.48)	-0.03	***
Radical innovation	0.25 (0.44)	0.28 (0.44)	-0.02	***
Product innovation	0.44 (0.50)	0.48 (0.50)	-0.04	***
Process innovation	0.45 (0.50)	0.50 (0.50)	-0.05	***
Open innovation	0.61 (0.49)	0.61 (0.49)	-0.01	NS
Neither	0.11 (0.31)	0.11 (0.31)	0.00	NS
Foreign ownership	0.14 (0.35)	0.16 (0.36)	-0.02	***
Public funding	0.14 (0.35)	0.16 (0.36)	-0.02	***
Exports	0.40 (0.49)	0.43 (0.49)	-0.03	***
Lack of employees	0.37 (0.48)	0.37 (0.48)	0.00	NS
Sales per person	11.68 (1.20)	11.74 (1.13)	-0.06	***
Number of external knowledge sources	2.46 (2.86)	2.66 (2.90)	-0.20	***
Machine	0.15 (0.35)	0.17 (0.37)	-0.02	***
External knowledge acquisition	0.02 (0.14)	0.02 (0.15)	0.00	***
Trainings	0.10 (0.30)	0.12 (0.32)	-0.01	***
Market research	0.16 (0.37)	0.17 (0.38)	-0.01	***
Design	0.06 (0.23)	0.06 (0.24)	0.00	**
Number of observations	33,270	41,849		

Notes.Welch Two Sample t-test; NA values are excluded. Standard errors are presented in parentheses.
Source: authors' calculations from PITEC

Table 19

PSM results with based on industry

a) Manufacturing industry with increase in share of women as treatment variable

Treatment groups	Variable	Sample	Treated	Controls	Difference (ATT)	Std. Err.	Significance
Group 1	*Open Innovation* (t+3)	Unmatched	0.651	0.655	-0.003	0.009	NS
	ATT	Matched	0.651	0.666	-0.015	0.011	*
Group 2	*Open Innovation* (t+3)	Unmatched	0.666	0.655	0.011	0.009	NS
	ATT	Matched	0.666	0.625	0.001	0.011	NS
Group 3	*Open Innovation* (t+3)	Unmatched	0.643	0.655	-0.012	0.012	NS
	ATT	Matched	0.643	0.665	-0.022	0.014	***

Notes. The number of observations in Group 1 is 10,000, in Group 2 it is 9,131 and in Group 3 it is 6,958.
Source: authors' calculations from PITEC

b) Manufacturing with increased in Blau index as treatment variable

Treatment groups	Variable	Sample	Treated	Controls	Difference (ATT)	Std. Err.	Significance
Group 1	*Open Innovation* (t+3)	Unmatched	0.649	0.655	-0.007	0.009	NS
	ATT	Matched	0.649	0.656	-0.010	0.011	NS
Group 2	*Open Innovation* (t+3)	Unmatched	0.664	0.655	0.009	0.010	NS
	ATT	Matched	0.664	0.669	-0.005	0.011	NS
Group 3	*Open Innovation* (t+3)	Unmatched	0.554	0.655	-0.101	0.018	***
	ATT	Matched	0.554	0.610	-0.056	0.022	***

Notes. The number of observations in Group 1 is 10,101, in Group 2 it is 9,309 and in Group 3 it is 5,312.
Source: authors' calculations from PITEC

c) Service with industry with increase in share of women as treatment variable

Treatment groups	Variable	Sample	Treated	Controls	Difference (ATT)	Std. Err.	Significance
Group 1	*Open Innovation* (t+3)	Unmatched	0.547	0.546	0.011	0.013	NS
	ATT	Matched	0.547	0.549	-0.002	0.015	NS
Group 2	*Open Innovation* (t+3)	Unmatched	0.562	0.546	0.016	0.014	NS
	ATT	Matched	0.562	0.547	0.014	0.016	NS
Group 3	*Open Innovation* (t+3)	Unmatched	0.513	0.546	-0.034	0.016	***
	ATT	Matched	0.513	0.543	-0.031	0.018	**

Notes. The number of observations in Group 1 is 6,044, in Group 2 it is 5,220 and in Group 3 it is 4,002.
Source: authors' calculations from PITEC

a) Service with increased in Blau index as treatment variable

Treatment groups	Variable	Sample	Treated	Controls	Difference (ATT)	Std. Err.	Significance
Group 1	*Open Innovation* (t+3)	Unmatched	0.552	0.546	0.007	0.013	NS
	ATT	Matched	0.552	0.556	-0.003	0.015	NS
Group 2	*Open Innovation* (t+3)	Unmatched	0.567	0.546	0.021	0.013	**
	ATT	Matched	0.567	0.553	0.01	0.015	NS
Group 3	*Open Innovation* (t+3)	Unmatched	0.461	0.546	-0.085	0.024	***
	ATT	Matched	0.461	0.505	-0.045	0.028	**

Notes. The number of observations in Group 1 is 6,048, in Group 2 it is 5,540 and in Group 3 it is 2,918.
Source: authors' calculations from PITEC

Table 21

Kernel for open innovation

a) Increased share of women treatment as treatment variable

Treatment groups	Variable	Sample	Treated	Controls	Difference (ATT)	Std. Err.	Significance
Group 1	*Open Innovation* (t+3)	Unmatched	0.609	0.617	-0.008	0.008	NS
	ATT	Matched	0.609	0.618	-0.009	0.008	NS
Group 2	*Open Innovation* (t+3)	Unmatched	0.625	0.617	0.009	0.008	NS
	ATT	Matched	0.625	0.623	0.003	0.008	NS
Group 3	*Open Innovation* (t+3)	Unmatched	0.593	0.617	-0.024	0.009	***
	ATT	Matched	0.593	0.613	-0.020	0.009	***

Notes. The number of observations in Group 1 is 17,138, in Group 2 it is 15,331 and in Group 3 it is 11,712.
Source: authors' calculations from PITEC

b) Increased Blau index as treatment variable

Treatment groups	Variable	Sample	Treated	Controls	Difference (ATT)	Std. Err.	Significance
Group 1	*Open Innovation* (t+3)	Unmatched	0.610	0.617	-0.007	0.007	NS
	ATT	Matched	0.610	0.619	-0.009	0.008	NS
Group 2	*Open Innovation* (t+3)	Unmatched	0.626	0.617	0.009	0.008	NS
	ATT	Matched	0.626	0.625	0.001	0.008	NS
Group 3	*Open Innovation* (t+3)	Unmatched	0.514	0.617	-0.102	0.014	***
	ATT	Matched	0.514	0.570	-0.056	0.015	***

Notes. The number of observations in Group 1 is 17,246, in Group 2 it is 15,834 and in Group 3 it is 8,785.
Source: authors' calculations from PITEC

Resümee

Sugudevaheline mitmekesisus töötajaskonnas ettevõtetes ja innovatsiooni tulemuslikkus: tõendid Hispaania ettevõtete tasandi andmete põhjal

Sageli leitakse teadustöödes, et soolisel mitmekesisusel on ettevõtete tegevustulemustele positiivne mõju. Samas on uuringuid soolise mitmekesisuse ning innovatsiooni sisendite või väljundite vahelise seose kohta suhteliselt vähe. Käesolevas artiklis uuritakse, kas sooline mitmekesisus ettevõtetes soodustab toote- ja protsessiinnovatsiooni. Täiendusena varasemale kirjandusele selgitatakse seost soolise mitmekesisuse ja avatud innovatsiooni vahel, st seost välise teadmuse allikate ettevõtte enda T&A-ga kombineerimisega. Empiiriline analüüs viidi läbi Hispaania ettevõttetasandi paneelandmebaasi PITEC andmete alusel, aastatest 2004–2016, kasutades tõenäosusliku sobitamise lähenemist. Leiame, et ettevõtted, kus naiste osatähtsus kogu tööjõus kasvab alla 10%, tegelevad tõenäolisemalt toodete, protsesside uuendustega ja avatud innovatsiooniga. Seevastu täheldame negatiivset seost avatud innovatsiooniga kui naiste osakaal kasvab ettevõttes üle 10%. Olulise tulemusena leiame, et soolise mitmekesisuse vähenemine on negatiivses seoses avatud innovatsiooniga ettevõttes. Meie tulemused toetavad mittelineaarse seose olemasolu.

Non-exclusive licence to reproduce the thesis and make the thesis public

We, Aruzhan Taufikova and Gulustan Ismayilova,

1. grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright, our thesis

GENDER DIVERSITY OF WORKFORCE AT FIRMS AND INNOVATION PERFORMANCE: EVIDENCE FROM FIRM-LEVEL DATA IN SPAIN,

supervised by Priit Vahter.

2. We grant the University of Tartu a permit to make the thesis specified in point 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 4.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. We are aware of the fact that the author retains the rights specified in points 1 and 2.

4. We confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Aruzhan Taufikova, Gulustan Ismayilova
18.05.2023