ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Bachelor Thesis [programme IBEB]

Title thesis: Can Gender Diversity Help Firms to Innovate? A Study of Association Between Gender Diversity in R&D Teams and Patent Applications of Business Enterprises

Name student: Diana Tretjakova
Student ID number: 596096

Supervisor: Elisa de Weerd
Second assessor:

Date final version: 28/06/2024

Table of Contents

Abstract

As many business enterprises continue to encourage diversity by introducing diversity quotas or making programs targeted towards female recruitment, and existing biases against female workforce participation are diminishing, it is valuable to examine to what extent this trend has been beneficial. This study thus aims to examine innovation, particularly patent applications of business enterprises, as innovation continues fueling competitive advantage and keeping business enterprises relevant and profitable. To perform the analysis, the study addresses the central research question of "To what extent is higher gender diversity in research and development teams associated with increased innovative patenting within business enterprises?".

The thesis utilizes a data set from multiple OECD and non-OECD countries to investigate the association between gender diversity and innovation outcomes. Gender diversity in R&D is negatively associated with the number of patent applications. The optimal level of BIH that maximizes the number of patent applications falls in the range of 0.300-0.350, ranging on the scale of 0 to 1. Thus, the analysis suggested a non-linear relationship between diversity and innovative patenting, as the number of patent applications tends to decline when further diverging from such BIH range. Moreover, higher gender diversity may be positively associated with patenting in specific categories, particularly of fixed construction, mechanical engineering, and electricity. Nevertheless, this finding remains open for interpretation considering the insignificance of coefficients obtained through the analysis.

The results provide some suggestions for business enterprises in terms of diversifying their R&D teams whilst reaching successful performance in innovation. This includes investing in effective diversity and conflict management and preventing extremely unbalanced teams, while actively fostering an inclusive culture through targeted programs and incentives. Like many previously conducted studies, the thesis emphasizes the complex relationship between gender diversity in the workplace and performance outcomes of business enterprises, yet contributes to the discourse by presenting a more specific analysis that hyper-focuses on R&D teams and innovative patenting.

Introduction

Innovation is an essential element of many firms, enabling them to gain a competitive advantage and differentiate themselves through unique products and positioning. The gradual improvement of processes or products of business enterprises not only ensures long-term profitability, but also may allow for more funds for the expansion and diversification of a business. Hence, innovative activities are at the core of many firms, particularly in high technology and rapidly changing industries (Ionescu & Dumitru, 2015). Moreover, innovation is not only desirable but oftentimes necessary for firms operating within countries with stricter rules of law in terms of privacy and environmental regulations. These may push certain firms to adjust their operational activities and ways of working, with innovation being the least costly option. A direct measure of a firm's innovation is patents, granting full, or temporary rights to a certain manufacturer over a product.

It is therefore crucial for businesses to find ways in which innovation, along with patents, can be maximized. The most basic way to examine this process is by looking at the base of those responsible for innovation within firms, which are largely its R&D teams. A firm's R&D team is the base component of human capital responsible for innovation and ensuring that the business keeps moving in an appropriate direction, rather than being stagnant. Hence, the composition of the R&D team is a fundamental solution to diversifying human capital and increasing creativity (Brennan, Ernst, Katz & Roth, 2020). Diversity provides a larger base of educational and personal backgrounds, which ultimately ensures higher representation. Such representation may generate ideas that teams contained to homogenous backgrounds may be less likely to come to, as individual creativity fosters collective creativity (Gong, Kim, Zu & Lee, 2012). Communication and ideas exchanges are both valuable pillars of innovation, as innovative processes are not static and limited to certain groups or geography of individuals and businesses (Jones, Chase & Wright, 2020). Diversifying human capital thus enlarges the possibility of building upon these pillars and maximizing the propensity to innovate.

In the recent decade, many businesses have already made efforts to incorporate gender diversity into their practices. Some have introduced mandatory gender quotas, while others have drafted non-discriminatory laws and policies, many of which continue holding support on a higher juridical level like in the EU (Belova & Ivanova, 2022). As a result, the number of women in business enterprises and other labor areas is growing and is at the largest it has yet been (Mazza, Furlotti, Medioli & Tibiletti, 2023). Nevertheless, women's participation, particularly in more science-focused and developmental fields is lagging in comparison to men's (EIGE, 2024). Even in 2019, only 20% of inventors worldwide were female, according to the recent OECD report on gender equality (2023). While the share of patents with at least one female inventor has more than doubled between 2019-

2013 (OECD, 2023), a vast majority of patenting applications are still contributed to by men. Consequently, most R&D teams in business enterprises are also dominated by men and team formation lacks gender diversity.

Research Contribution

This research will help contribute to understanding whether such a structure is optimal for firms' progress and innovation, or if gender diversity is positively associated with patenting. The potential association will be tested by comparing the number of patent applications and gender diversity levels in the R&D teams of 21 OECD and non-OECD countries, through the use of panel data of these countries for the years 2000-2020. If a positive link is found between innovation and gender diversity in R&D teams, it can act as an indicator for business enterprises to change their current business practices or offer strategic implications for their human capital. Considering that innovation is one of the major pillars of successful and profitable businesses, the role of human capital must be considered to achieve the optimal outcomes for the long-term survival of business enterprises.

While the impact of gender diversity on various performance factors and outcomes of business enterprises has been studied in relatively great amounts, most studies, such as those by Simionescu, Gherghina, Tawil, and Sheikha (2021) or Gruszczynski (2020), have focused on gender diversity in enterprise ownership or board membership. Contrastingly, a significantly lower number of studies examine the relationship between gender diversity and performance measures of smaller individual teams, such as R&D teams, within business enterprises. Nevertheless, R&D teams have been an integral part of companies' employment, particularly following the technological boom and increasing levels of globalization (Mowery & Rosenberg, 1998).

Despite R&D teams' direct contribution to innovation, innovation itself is difficult to measure, as recognized in research performed in the sector. Keith (2009), for instance, recognizes that current innovation metrics all have their strengths and weaknesses. For this reason, current research has utilized a wide range of innovation measures, acknowledging their shortcomings. This research will aid in filling the current literature gap in the link between R&D teams and innovation, particularly by examining the number of patent applications. By using the number of patent applications as an innovation measure, it aims to help direct business enterprise policy regarding gender diversity in business enterprise teams like those of R&D. Hence, the research aspires the results may provide strategic implications for business enterprises regarding optimal team formation and ways in which gender diversity can be utilized to maximize innovative outputs.

The research will begin with presenting the main research question and sub-questions. Afterward, the literature section will explore the already existing findings on the topic, followed by a data and methodology sections. Afterwards, the empirical results from the methodology will be presented. A conclusion including the discussion, limitations, and suggestions for further research derived from the analysis will follow.

Main Research Question

To achieve evaluation of the association between gender diversity and innovation outcomes, the thesis will explore the role of gender diversity in research and development (R&D) teams in business enterprises, particularly, the participation of women in R&D teams. Gender diversity will be explored based on its association with the innovative activity of business enterprises as an outcome. Innovative activity is quite a broad term that can be interpreted in various aspects, as even in the sole sector of business enterprises, innovation can be measured and understood differently. The research of this paper will concentrate on measuring innovation in terms of patent application by business enterprises in different countries. Hence, the main research question is:

“To what extent is higher gender diversity in research and development teams associated with increased innovative patenting within business enterprises?”.

Sub Questions

In addition to the central research question, the paper will aim to answer two additional sub-questions to explore the topic of gender diversity in R&D teams in more depth. The sub-questions will aid in providing more concrete suggestions for business enterprises in terms of their hiring practices and team formation. The first research question will allow to make more concrete conclusions about generalization of findings and whether they can be applicable to all patent applications, or rather just certain types of them. Different patent applications can be categorized into varying patent categories, however, the main research question pools all categories together. Hence, exploring whether gender diversity is associated with increased patenting in only some patent categories can help to provide a more nuanced answer to the main research question. The OECD database web page currently has data on 8 different categories of patents, ranging from human necessities to electricity, and their distinction will be further discussed in the Data and Methodology section. If changes only in specific patent categories are related to gender diversity, this would impose different implications for business enterprises in various industries, posing the sub question of:

“What type of patents are most associated with changes in gender diversity in R&D teams?

Moreover, some authors of previous literature point to the possibility of observing a non-linear effect between gender diversity and innovation outcomes. Several studies, such as those conducted by Kou et.al. (2019) or Diaz-Garcia et.al. (2013), recognize that certain levels of gender diversity may have the potential to be detrimental to a firm's innovation outputs. Thus, a question that would aid in pinpointing a potentially optimal level of gender diversity that is able to yield the largest additional increase in patent applications, can reveal certain conclusions regarding the direction of the association between gender diversity and innovative outputs. Moreover, it can help to identify a level at which increases in gender diversity may have a negative association with the number of patent applications, if any such threshold exists. The upcoming underlying theory and literature discussion further explores the concept of varying implications of gender diversity on patent applications, acting as supporting precursors for the second sub-question:

"What level of the Blau index of heterogeneity in R&D teams yields the highest number of patent applications filed?"

Underlying Theory and Literature

Proponents of Gender Diversity in R&D Teams

The existing literature yields varying perspectives on the impact of gender diversity of R&D teams on patenting and other innovative metrics of firms. Most research largely favors diversity in innovative activity to some degree, providing evidence for its benefits with regards to a contribution to a firm's innovative outputs. Martinez et.al. (2017), for instance, studies the impact of gender diversity in R&D teams on radical innovations by examining activity in various Spanish firms, finding a positive and significant relationship between gender diversity in R&D teams and the propensity of firms to produce radical innovations. They attribute these results to gender diversity increasing interpersonal skills and promoting better work dynamics. Similar results of gender diversity in R&D teams positively impacting innovative performance are also observed by Turner (2009), who suggests that diversity within teams can help to navigate more complex tasks while generating groundbreaking solutions both in teams and individuals themselves. Both such by-products of gender diversity are therefore able to contribute to higher rates of innovative patenting in firms.

Capozza and Divella (2023) even go as far as concluding that workforce diversity in all aspects of European firms' capital increases R&D decisions and intensity, positively affecting both product and process innovations. They explore gender diversity in different levels of the workforce and gender diversity in ownership, whilst analyzing their effects on firms' productivity which includes innovative development. The authors stress that this success is the outcome of learning and performance, which diversity can foster through interaction and exchange of knowledge, improving creativity and problem solving in the process.

While siding with the consensus that gender diversity in R&D teams has some association with innovative activity of firms, certain studies conclude that this does not apply to all kinds of innovations. Teruel and Segara-Blasco (2017) utilize data from the Spanish Community Innovation Survey containing panel data on differing Spanish enterprises, finding that the impact of gender diversity varies with innovation type. The large positive impact was observed comparatively more for non-technological, such as marketing or organizational, rather than technological innovations, being product or process innovations.

Gender Diversity in the Workforce

Many of the studies that support a positive association or correlation between gender diversity in firms and innovative activity, however, do not specifically focus on gender diversity within R&D teams, but rather explore gender diversity as a total of the firms' employment, or gender diversity of board membership. This is due to the lack of exploration of gender diversity particularly in R&D teams, as many studies focus on female management and ownership as a decisive factor for firms' performance metrics, with more niche parts of business enterprise networks, such as R&D teams, being an area less investigated. For instance, the benefits of female employment for innovative activity are supported by a study of Pfeifer and Wagner (2013) on increases in female employment in German manufacturing firms. Although the authors examine female employment in the firm on a whole level, not only in R&D teams, but they also find a significant positive correlation between it and R&D activity. Hence, they propose higher female participation in the workforce may even be used as an instrument to counteract the negative effects of an ageing workforce on innovation.
Successful impacts of women's employment in the workforce has also been supported by studies regarding green processes and product innovations. Nadeem, Bahadar, Gull and Iqbal (2020) find that board gender diversity has a significant positive association with environmental innovation, in the context of United States-based enterprises. Green process innovation and green patenting are also tested by He and Jiang (2019), who use panel data from Chinese manufacturing firms to prove their claims. The authors conclude a systematic relationship between higher female board representation and the occurrence of firm-level green innovation.

Psychological Evidence

The mechanisms for greater innovative propensity of mixed gender R&D teams have also been justified by psychological theories. Kou et.al. (2019) utilize the cognitive resource diversity theory to explain how gender diversity promotes process innovation and yields greater efficiency. The theory states that groups composed of diversified members will likely yield better performance outcomes due to unique combinations of cognitive perspectives and resources which diverse members can bring. Kou et.al. (2019) utilize panel data from regional Chinese research institutes, investigating the association between gender structure in these R&D teams on different kinds of R&D output, including patents. They find evidence that gender structure has both positive and negative implications for R&D, depending on the output, and that female and male personnel are better at different kinds of research.

Gonzalez-Moreno et.al. (2018) propose a similar concept of social diversity perspective, which largely points to the idea that higher diversity enhances the quality of a decision through greater information sharing between agents and broader perspectives. Such mechanisms create intrinsic value in diversity, which can ultimately impact different aspects of firms' activity, such as revenue, customer base, profits, and innovation. Gonzalez-Moreno et.al. (2018) use this concept to explain the positive link between mixed gender teams and product innovation of national and international manufacturing firms. The findings confirm that gender diversity in R&D teams can help firms to maintain a competitive advantage in their products, processes, and operations, by aiding in innovation.

Wikhamn and Wikhamn (2019), on the other, propose a different theory that diverts from the favorable view of gender diversity and its contribution to positive factors such as productivity, creativity, and innovation. While exploring innovative performance in Swedish firms depending on the firm's gender diversity in R&D teams, they consider the social categorization perspective as a potential deterrent to the success of gender diversity. The theory argues that individuals tend to classify themselves in groups based on certain characteristics, like gender, thus comparing themselves to others within and outside the perceived group. Such abstract categorization may cause individuals to perceive those within their group as more trustworthy, engaging, and desirable for the exchange of information than those outside their group (Paulus & Njstad, 2019). Hence, adding more gender diversity into a team will create barriers to the characteristics desirable to achieve innovative outcomes, such as those of creativity, and constant idea exchanges. Until the in-group remains sufficiently larger than the out-group, female workers are unlikely to experience the workplace as more open and engaging, when being a minority in it. This would likely start to change only after the out-group grows sufficiently so that the in-group is no longer large enough to limit interaction with the out-group. If this interaction grows, the mechanisms of the social categorization theory weaken, and idea exchanges will be facilitated by a need to interact with a diverse team, fostering innovative outcomes in the process.

Issues Associated with Gender Diversity

Nevertheless, many studies point out the potential negative implications of gender diversity, and the great impact of the context of the industry. Kou et.al. (2019) argue that diverse groups are more prone to conflict and internal tensions, which can detrimentally impact group performance as a whole, if the tensions are persistent and not resolved for prolonged time periods. They partly use this reasoning to justify their findings of a higher proportion of female R&D personnel in regional Chinese research institutes being associated with lower R&D efficiency. This caveat is also recognized by Diaz-Garcia et.al. (2013), who consider diversity to be a “two edged sword” (pp. 149). They emphasise that increased diversity creates a greater need for coordination, interaction and communication within a team, which can be difficult to impose in some workplaces. If not managed properly, such dissonances can actually hinder team performance and limit the incentive for knowledge sharing which is essential for patenting and innovation.

Moreover, Na and Shin (2019) cite higher risk aversion among women as a factor potentially hindering innovative activities in R&D teams with higher percentages of women. They examine women’s participation in firm ownership, roles in top management and female majority in the workforce in various emerging countries, finding varying effects of each category on different innovation metrics such as marketing, product or process innovations. The authors also point out that country and workforce culture may play a more significant role in innovativeness than team gender diversity, hence downgrading its role in fostering greater innovation.

Some studies therefore propose that diversity and innovation may have a nonlinear relationship with one another, hence indicating that more gender diverse teams are only beneficial until a certain threshold. Capozza and Divella (2023) do find a non-linear relationship between a firm’s productivity and gender diversity in the workforce, concluding that increasing women’s participation beyond a certain threshold of under a 50/50 ratio of male to female, may be detrimental for performance. Such non-linear U-shaped relationship is also observed by Gonzalez-Moreno et.al. (2018), who utilize the social cognitive theory to explain their suggestions to companies to encourage gender diversity in R&D teams in manufacturing firms only up to a certain optimal threshold. The authors attribute this to interpersonal conflicts associated with higher gender diversity, jeopardizing cooperation and effective communication, and reducing innovative capacity of teams in the process. According to Gonzalez-Moreno et.al. (2018), higher levels of gender heterogeneity fuel intergroup bias and social categorization, both hindering innovation. Hence, more homogenous teams with a ratio of male to female being lower than a 50/50 split would be more favorable for the teams’ performance outcomes.

The duality of the nature of the relationship between gender diversity and innovation is similarly suggested by Liu, Ozer and Zhou (2023). In their research, the authors test for a curvilinear relationship between different kinds of diversity, such as gender, age, ethnicity, and innovative performance of R&D teams. Low to medium degrees of diversity were found to have a positive relationship with team integration, while medium to high degrees to have a negative relationship with team integration, therefore rather proposing lower levels of diversity to be more optimal for innovative output. Like Gonzalez-Moreno et.al. (2018), the study ultimately concludes the presence of a U-shaped relationship between the two variables. Such diminishing benefits of diversity are also presented in the study of the largest energy companies in Brazil, conducted by Neves Cavazotte and de Oliviera Paula (2020). Although the authors do not solely focus on gender diversity, but rather on shared and diverse leadership within firm's R&D teams, they find shared leadership positively affects creativity and absorptive capacity of R&D teams only to a certain level. Instead, the relationship is curvilinear, particularly U-shaped, as higher levels of shared leadership result in synergy losses, preventing R&D teams from operating effectively and communicating successfully.

Such studies which find potential detrimental effects of gender diversity both in the workforce and in R&D teams specifically, yield contrasting results to the papers that find beneficial components of gender diversity and its association with innovative activity. Additionally, the existing psychological theories and evidence are able to provide explanations both for the negative and positive sides of having a more gender diverse team in the workplace, making it difficult to draw concrete predictions and hypotheses regarding the results of this research. This is especially valuable to take into account, considering the inter-country nature of the research. In certain parts of the world, including developed countries, women still face great obstacles in obtaining higher education and participating in male dominated positions. Additionally, the role of gender diversity may be non-linear because it is not the only factor impacting R&D team performance. While this research will control for multiple variables which could potentially influence patent applications, it cannot account for many country-specific measures, or outside factors which influence the success of R&D teams.

Hence, due to the conflicting evidence regarding the relationship between gender diversity in R&D and the number of patent applications, a study of explorative nature will be performed. Considering the varying perspectives in the literature, the direction of the relationship between the variables remains difficult to grasp. A two-sided test will therefore be used, with the Null Hypothesis being that there is no relationship between gender diversity in R&D personnel and the number of patent applications. Meanwhile, the Alternative Hypothesis will be that there is a relationship between the two variables. By proposing a rather open hypothesis, the research will focus on exploring the relationship all together, rather than hypothesizing a specific direction of it.

Data

The research will use macro level country-data taken from the OECD Statistics database website (OECD Statistics, n.d.), which has data on both OECD and some non-OECD countries. The OECD Statistics database holds data regarding various macro-economic aspects on countries, such as industries, consumers, lifestyles, innovation and digitalization. Particularly, data from the reports on science, technology and innovation policies will be used and data specific subsections will be taken, depending on the variable to be explored.

While using macro level data instead of micro level data does yield a lower number of observations, the chosen data set provides data on the variables needed to explore the research question. The dataset incorporates not only data on the dependent and independent variables, but also data of potential control variables. This dataset is available on a macro level, hence, macro level data was used in the research. Moreover, the data source of OECD Statistics is a highly reliable source of information, being one of the world's largest macro-level data sources. The OECD is an international organization which cooperates with smaller inter-country organizations, allowing it to gather up to date information from sources others may not have access to. Moreover, it is involved in individual country-specific policy drafting, further validating the reliability of its data (OECD Statistics, n.d.).

In order to keep a homogeneous and comparable sample, all data will be taken from the business enterprise sector. Certain industries and sectors have a higher propensity to innovate, particularly high technology sectors, hence it would be unreasonable to gather and compare data from all sectors at once. The data will be taken from years 2000-2020, as these are the years for which the data is available for the combined dependent, independent and control variables. Moreover, the time period allows to gather a sufficient number of observations to draw comparisons using panel data. As some countries from the database have missing years for certain data points, only countries with a maximum of eight years of missing data – and thus a minimum of twelve years observed - will be investigated in the research being both OECD and non-OECD economies to provide more observation. Such selection criteria results in a total of combined 53 country-year observations from 5 countries being dropped. This totals to 21 countries (6 non-OECD and 15 OECD) to be examined, yielding 315 country-year observations.[1]

The dependent variable to be examined is the total number of patent applications filed from innovators, per country of residency. Filed patent applications will be regarded as a direct measure of innovation. Patent applications are an informative measure of innovation, as they grant their

[1] The OECD countries used in the research are: Belgium, Chile, Czechia, Estonia, Hungary, Ireland, Italy, Latvia, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, and Turkey. The non-OECD countries are: Argentina, Taiwan, Romania, Russia, and South Africa.

owners the rights over a certain product or process for an (in)definite amount of time. Hence, patents are created to ensure that such innovations cannot be replicated by others, due to their superiority and newness. Thus, patents are a direct measure of innovation which is also simple to measure. Patent applications are binary numbers, which are public knowledge and can be counted for business enterprises. Rogers (1998), notes that patent applications are a valid proxy measure of innovative output, as even in cases patent applications get rejected, they nevertheless indicate innovative behaviour by firms. Acs, Anselin, and Varga (2002) use regression analysis to investigate whether patents are an appropriate measure of production of new knowledge. They conclude that patents provide a reliable measure of innovative activity, and "capture technological change" (Acs et.al., 2002, p. 1080). Similar studies which aim to test the reliability of using patents as innovation metrics, such as those conducted by Katila (2000) and Nagaoka, Motohashi and Goto (2010), deduce alike findings, strengthening the methodology of the research of using the number of patent applications as an output measure of innovation in firms.

Despite being a broad concept that can be measured, innovation can be split into product and process innovation. Product innovation incorporates introducing a new product or even service to the market, or a product with significant improvement from its previous version. Meanwhile process innovation relates to new elements introduced to restructure a specific way of doing or manufacturing (Nandal, Kataria & Dhinga, 2020). Patent applications are therefore a valuable way to measure innovation of a business enterprise, as they can incorporate both process and product innovations combined. Generally, public data sources do not have separate data and applications for process and product innovation. The differentiation between the two is therefore not possible and cannot be explored in this research. However, keeping in mind that both are incorporated into patents is useful when making potential conclusions and recommendations for business enterprises and their R&D teams.

The independent variable to be examined is the share of internal R&D personnel of the country, that is female. It is significant to consider, for the purpose of drawing conclusions in this research, that the share of female workers in R&D may not necessarily reflect the share of female workers in management positions. Hence, the research will avoid drawing conclusions or connections to female top management and ownership, based on the findings. The share of female workers in R&D will be transferred in accordance with the Blau index of heterogeneity (referred to as BIH later in the paper), which is used to quantify the level of diversity within a group. Thus, only the Blau index will be used as an independent variable and is a direct reflection of the share of R&D personnel by each country. The BIH ranges from 0 to 1, with 1 indicating most diversity and 0 the least. Using this index would allow for more clear comparisons and better quantification of diversity within R&D teams. Moreover, the BIH has been a widely used index in diversity-related studies, which have also focused on innovation and gender diversity within teams. For instance, studies by Diaz-Garcia et.al. (2013), Teruel and Segarra-Blasco (2017) and Capozza and Divella (2023), have all used the BIH for its elaborateness in comparison to a simple ratio measure, considering that the index represents diversity in terms of proportions. It quantifies diversity by measuring the probability that two members randomly selected from a group would be from different categories, hence, in the case of this research, if the members would be of different gender.

The BIH is calculated as follows, with p representing the proportion of each category in a population, meaning the proportion of women in comparison to men in this research:

$$BIH = 1 - \sum (p_i)^2$$

To account for contextual variables and control for those which might influence the number of patent applications per country and also the share of female workers in R&D, several control variables will be introduced. The first is the gross domestic expenditure on R&D, which as it increases, will likely increase patent applications. Countries that have larger R&D expenditures are also generally more likely to be developed and wealthy, with both factors being associated with more gender equality and potentially higher women's representation in R&D. Moreover, Czarnitzki and Hottenrott (2011) explain that R&D investment is a critical driver of innovation in firms, leading to increased patenting capacity. Second is the government budget allocation for R&D, which would also positively impact patent applications filed. This variable similarly relates to diversity within R&D teams as the first variable, likely being higher as the women's share in R&D teams increases. R&D investments are drivers of innovation, with government spending in particular being a large part of such investments (Hall, et. al., 2010). Moreover, Guellec and van Pottelsberghe de la Potterie (2003) find that government R&D funding positively impacts business R&D investment and thus increases innovative outputs, further emphasizing the need to control for this variable.

The third control variable is the R&D tax expenditure and direct government funding of R&D, which by prediction should increase the dependent variable. The motivation behind this control variable is similar to that discussed for the government budget allocation for R&D, with the same studies supporting its usage. Then, the research will control for business enterprise researchers as a percentage of the national total labor force involved in research and development, along with the total R&D personnel per thousand of total employment. It is predicted that both variables will positively impact filed applications for patents, as higher R&D personnel would result in more input and knowledge sharing. This has been supported by Hall, Lotti and Mairesse (2009), who conclude that intensity of R&D activity which is influenced by the size of R&D personnel, is correlated with innovation outcomes like patents. However, quality of R&D personnel may not necessarily equate quality, so the direction of effect of this control variable is rather ambiguous.

While the utilized control variables all impact patent applications, some may also be outcomes of them. For instance, more patent applications would indicate that business enterprises have the potential to be successful and innovative, which in turn would lead to higher future R&D expenditure and higher gross domestic expenditure on R&D. While this direction of the relationship is less likely, it cannot be ruled out. Therefore, the study cannot confidently claim a purely one-directional relationship between the control variable and the dependent variables. However, these are the best available control variables considering the dataset and the fact that sensitive micro-level data is rarely disclosed by business enterprises. It is difficult to assess the relationship without running an experiment or having micro-level data, both of which are not viable for the nature of this research. This discourse leads to a possibility of reverse causality, which is further discussed in the section on Limitations of the Research.

There are also some control variables that would greatly benefit the research, but are unavailable. These include subsidies and grants towards R&D, both from the government and private individuals or institutions, as those can be particularly helpful towards increasing patents by providing additional resources to innovators and attracting more female personnel. Moreover, the enforcement of intellectual property rights would be valuable to control for, as not all countries respect patent rights equally. If law enforcements do not offer strong protection of property rights, this disincentivizes patent applications. Controlling for the state of technological development would also be preferrable, as majority of patenting innovations require access to technological instruments and research facilities. Notably less technologically developed countries would be unable to produce many patents, simply due to a lack of facilities and instruments required for appropriate research.

Descriptive Statistics

Descriptive statistics of each variable used for the research were constructed to obtain initial information about the data. As observed in the descriptive statistics of Appendix A, gender diversity is in general not high in R&D, with the BIH ranging from 0.125 to only 0.5 out of 1. Hence, the probability that two members randomly selected from a group would be in different categories is no more than 0.5. Additionally, the number of patent applications greatly varies by country, having a standard deviation of 212.652, indicating that countries yield a very varying amount of patent applications. Large standard deviations can be observed in all variables of Appendix A, further confirming the differing policies towards R&D in business enterprises and the personnel involved in R&D. There are thus considerable differences between countries in their outcomes and approaches to patents and R&D, emphasizing the importance of controlling for the chosen control variables.

Additionally, a table ranking countries based on all variables used in the research was constructed. It includes three countries with the smallest values of each variable and three countries with the largest values of each variable. This was done by averaging all observations per variable per country, and then ranking these averages. Appendix B showcases that 2 out of 3 most gender diverse countries with the highest BIH are non-OECD countries of Russia and Romania. Meanwhile, the other most R&D gender-diverse country is Latvia, which geographically is close to Russia and Romania. While Latvia is geographically a Northern European country, its cultural and historical heritage is close to Eastern European. Hence, the most gender diverse countries are those in the Eastern European and Northern European regions. Regarding the least gender diverse countries, Central European Hungary and Czechia are at the top, also with a non-OECD country of Taiwan.

However, when considering the number of patent applications, none of the most gender-diverse countries have the highest applications. In fact, the highest number of patent applications is in Taiwan, which is also the least gender diverse. Meanwhile, Latvia, while being the most gender diverse, also has the lowest number of patent applications. Nevertheless, this may be explained by Taiwan having the highest gross domestic expenditure on R&D, business enterprise researchers as percentage of national total, government budget allocation for R&D and total R&D personnel per 1000 of employment. Latvia on the other hand ranks as one of the lowest in these categories. Other than the above, no sticking pattern regarding the BIH, number of patent applications, and other control variables can be observed.

Methodology

First, a simple ordinary least squares regression (OLS) will be run to execute the methodology. This regression will only include the dependent and independent variables and will serve as a starting point for the analysis. It will also enable better comparison of other models that will follow the OLS. Thus, for a more complex analysis, a multiple linear regression for the panel will then be run on the dependent variable, including the controls. This will provide a more accurate model for the data set and control for variables that can impact the results. The following regression will be run:

$$
\begin{aligned}
\text{y (\# of patent applications)} = {} & \beta_0 \\
& + \beta_1 x(\text{Blau diversity index})_{ct} + \beta_2(\text{gross domestic expenditure on R\&D})_{ct} \\
& + \beta_3(\text{government budget allocation for R\&D})_{ct} \\
& + \beta_4(\text{R\&D tax expenditure and direct government funding of R\&D})_{ct} \\
& + \beta_5(\text{business enterprise researchers as a percentage of the national total})_{ct} \\
& + \beta_6(\text{total R\&D personnel per thousand total employment})_{ct} + e_{ct}
\end{aligned}
$$

In the model, y stands for the number of patent applications filed by a country in a specific year from 2000 to 2020, hence being the dependent variable in the regression. The constant is represented by β_0, while the Blau diversity index variable, which measures the BIH per country per specific year is captured by the coefficient of β_1, with x indicating the interaction term of the independent variable. The rest of the constants of $\beta_2 - \beta_6$ represent the constant for the appropriate control variable. The error term of the regression is captured by e_{ct}. The subscripts c and t stand for 'country' and 'time' respectively. Since panel data is used, each observation is given in a specific country, in a certain year from 2000 to 2020 timeframe. This model is a relatively simple indicator; however, it is fitting considering that the dependent variable is non-binary and continuous. Such a model will allow an understanding of the magnitude and direction of the relationship.

To determine whether the selected control variables are appropriate to use and have an impact on the independent and dependent variables, each control variable will be individually added to the model. The coefficient change of the Blau index of heterogeneity will then be evaluated as each additional control variable gets added to the model. A change, particularly if a significant one, will indicate that the selected control variable is likely appropriate and has potential to affect the results of the model, hence, needing to be controlled for. However, this line of though may potentially not be true in the case of reverse causality. Reverse causality is a limiting factor of this research and is discussed further in the Limitations section. Despite its potential to impact the results, in this research, the outlined methodology is the most fitting way to determine whether the chosen control variables are appropriate to use.

Considering that the data are panel data and changes in the dependent and independent variables are observed over time, in theory both fixed effects and random effects are possible to be applicable to the model. Utilizing fixed and random effects will allow the methodology to control for either time-invariant or time-variant characteristics respectively. This is much needed for panel data, as observations are bound to vary over time, with the variation being attributed to certain factors. Thus, fixed or random effects will help to mitigate the impact of omitted variables bias, as some time variant characteristics may not be controlled for. As research utilizes panel data, investigating whether random or fixed effects are the most fitting will allow for the determination of the nature and assumptions of underlying individual-specific effects. Accounting properly for individual-specific effects and unobserved heterogeneity in data allows to determine the model most suitable for answering the research question. To figure out whether a model with fixed effects or random effects should be used, a Hausman test will be performed. First, the proposed linear regression model will be estimated by applying fixed effects and then by applying random effects.

In a similar manner, the proposed model will be tested for potential autocorrelation, which could indicate correlation between panel data observations at different time periods, with such observations being non-independent from one another. The presence of autocorrelation can indicate valuable information about the data, having various adverse effects on the anticipated regression analysis. Autocorrelation can lead to invalid predictions of the model, by violating the assumption of independence of errors, thus leading to biased coefficient estimates which inaccurately reflect the relationship between variables. If time periods where autocorrelation is not accounted for are present, autocorrelation can lead to less reliable predictions of the model, as errors across time periods are in some way related to one another. Consequently, linear regression estimation coefficients may be estimated incorrectly, leading to unfounded and incorrect conclusions and analysis of the results.

The above-mentioned methodology of the research rests upon the assumptions of multiple linear regressions, such as normality or errors, no multicollinearity, and the independence of errors. Some of these assumptions cannot be tested and are imposed onto the research to proceed with the methodology. Moreover, the potential presence of autocorrelation and heteroskedasticity can undermine the reliability of the results, by violating the previously mentioned assumptions. In the case of their presence, omitted variable bias, sampling variation and non-linearity are issues of even greater concern. However, the research mitigates these concerns by aiming not to establish a causal relationship, but rather an association between the main variables of the research question. While this decreases the strengths of the analysis and may prevent from drawing strong and definitive conclusions, it prevents from drawing unfounded claims regarding a causal relationship. Moreover, it minimizes the concerns regarding violation of assumptions of multiple linear regression, whilst still not preventing the shortcomings of the methodology from yielding valuable insights in the analysis.

Sub Question 1

To address the first sub question of whether changes in patent applications are differently related to gender diversity in R&D teams, a multiple linear regression will be performed in a similar manner as for the original research question. However, the regression was run 8 times, with each time for a separate patent category for the number of patent applications being the dependent variable. The data for each patent category was available on the same OECD webpage in the same science, technology, and patents subcategory. The patent categories are classified by alphabetical letters and their explanation can be found in Appendices E and F. The correlation coefficient and its magnitude for each patent category will then be examined to find which of the 8 patent categories are related to gender diversity, if any are related at all.

Sub Question 2

Next, the second sub question focusing on investigating the level of the Blau index of heterogeneity leading to the highest increase in the number of patent applications, will be investigated by introducing dummy variables to the regression equation. Dummy variables for 6 different Blau diversity index categories will be introduced, namely 0-0.250; 0.250-0.300; 0.300-0.350; 0.350-0.40-; 0.400-0.450; 0.450-0.500. The ranges were selected based on the lowest and highest Blau diversity index values in the dataset, being 0.215 and 0.500. Each dummy will be examined in terms of the increase (or decrease) it yields for the number of patent applications. The dummy category which yields the highest increase in the number of patent applications will be regarded as the level of gender diversity which maximizes the dependent variable, and it thus optimal for firms.

$$
\begin{aligned}
& y\ (\#\ of\ patent\ applications) \\
&= \beta_0 + \beta_1(\text{BIH1})_{ct} + \beta_2(\text{BIH2})_{ct} + \beta_3(\text{BIH3})_{ct} + \beta_4(\text{BIH4})_{ct} \\
&+ \beta_5(\text{BIH5})_{ct} + \beta_6(\text{gross domestic expenditure on R\&D})_{ct} \\
&+ \beta_7(\text{government budget allocation for R\&D})_{ct} \\
&+ \beta_8(\text{R\&D tax expenditure and direct government funding of R\&D})_{ct} \\
&+ \beta_9(\text{business enterprise researchers as a percentage of the national total})_{ct} \\
&+ \beta_{10}(\text{total R\&D personnel per thousand total employment})_{ct} + e_{ct}
\end{aligned}
$$

Where BIH refers to the different BIH categories of BIH 1 0.250-0.300; BIH 2 0.300-0.350; BIH 3 0.350-0.400; BIH 4 0.400-0.450; BIH 5 0.450-0.500 and BIH of 0-0.250 is used as the reference category.

Additionally, to corroborate the answer to this sub question, and to determine whether it is in line with the previously mentioned literature, a scatterplot will be presented of a relationship between Blau index of heterogeneity and the number of patent applications. A parabolic regression will then be applied, by regressing the Blau index of heterogeneity on the number of patent applications. Based on the regression, the local maximum will be calculated, indicating an optimum level of gender diversity which would be most desired by firms which wish to maximize their innovative outputs of patent applications. The maximum point will be compared to the level of the Blau index of heterogeneity previously concluded to yield the highest increase in patent applications, preferably yielding similar results.

Empirical Results

Research Question Findings

To provide a full evaluation of a potential association between gender diversity and the number of patent applications, five different regressions models were performed. All models in Table 1 provide similar results, showcasing that business enterprises with a higher diversity index on average have a lower number of patent applications. In terms of BIH, this indicates that on average, as heterogeneity within R&D teams increases, hence becoming more gender diverse, the number of patent applications decreases. For instance, in model (3), an additional increase in a BIH unit is associated with a decrease of 373.723 patent applications. Similarly, in model (5), an additional increase in a BIH unit is associated with a decrease of 359.053 patent applications. This can be observed by the negative and significant coefficient of BIH, which is present in all models, either at a 0.05 or 0.01 significance level. The coefficient of BIH is almost twice as high for the model (1) in comparison to all other models. This is likely due to the nature of the data utilized for this research being panel data and due to omitted variable bias. Panel data requires controlling for certain time-varying effects which may either vary between countries or within countries throughout the years. Hence, not controlling for these effects by imploring random or fixed effects changes the coefficient of BIH. Meanwhile, not controlling for certain variables overestimates the BIH coefficient, by attributing variation solely to that coefficient.

Additionally, the coefficients of BIH become significant at a higher level of significance with the addition of control variables, changing the level of significance from 0.05 to 0.01. The BIH coefficients themselves do not differ by a large amount between the models, although are somewhat higher for models (3) and (5) which impose random effects. Meanwhile, the constant remains positive and significant at a 0.01 level in all models. The significant control variables in models (4) and (5) are general direct expenditure on R&D in a country and the total number of R&D personnel per 1000 of the labor force.

Table 1

Regression results of BIH onto the number of filed patent applications.

Variable	**OLS** **(1)**	**FE, General** **(2)**	**RE, General** **(3)**	**FE, Controls** **(4)**	**RE, Controls** **(5)**
BIH	-616.500***	-337.236**	-373.723**	-337.827***	-359.063***
	[178.010]	[153.460]	[151.958]	[124.943]	[123.813]
General direct expenditure				0.023***	0.022***
				[0.003]	[0.003]
Government budget				-4.720	-2.835
				[56.720]	[56.059]
% of total labor force				0.174	0.291
				[0.638]	[0.626]
Total per 1000 of labor force				10.188***	10.986***
				[3.315]	[3.269]
Tax and funding				-64.523	-91.371
				[154.835]	[153.031]
Constant	450.973***	342.916***	416.219***	169.793***	207.756***
	[69.808]	[59.560]	[82.005]	[59.067]	[72.672]
R Squared	0.037	0.195	0.195	0.565	0.571
Observations	315	315	315	249	249

Note: * indicates $p < 0.1$, ** indicates $p < 0.05$, *** indicates $p < 0.01$, standard errors are depicted in brackets. The regressions represent random and fixed effects time series regression models, with and without control variables, in OECD and non-OECD countries in the years 2000-2020.

To test for whether the random or fixed effects model is more appropriate for the analysis, a Hausmann test was performed (see Appendix C). The Null Hypothesis of the test states that the difference in coefficients of fixed and random effects is not systematic, while the alternative claims such different is in fact systematic. In the case of this analysis, with the Chi squared p value (0.089) being smaller than the selected level of significance of 0.1, with a 10% significance level chosen due to the relatively small sample size of the dataset. Hence, the Null Hypothesis is rejected, indicating that the fixed effects model is a better fit. A 10% significance level was chosen due to the relatively small sample size of the dataset. This rejection of the Null Hypothesis is evident considering the nature of the dataset and the research. The purpose of the research is to examine differences between years within each individual country, which the fixed effects model considers by assuming individual specific effects are uncorrelated with independent variables. On the other hand, the random effects model focuses on the differences between countries at a specific point at time, in the case of this research at a specific year between 2000 and 2020.

To test whether the chosen control variables were appropriate to use for the analysis, each control variable was added individually to the multiple regression (see Appendix D), as described in the Methodology. As observed in Appendix D, all control variables are appropriate to utilize for this research, as they change in the BIH coefficient. As observed in Appendix D, the BIH coefficient changes and remains significant at either 0.05 or 0.01 significance levels for all control variables. This may indicate their ability to influence the independent and dependent variables. While a significant change in the BIH coefficient does not necessarily mean the control variables are appropriate to use, it does further validate the reasoning behind choosing them. A striking and significant coefficient is that of the government budget, being negative, and indicating that on average, as the government budget for R&D activity increases, the number of patent applications decreases. This may be explained by the budget going towards other R&D activities that are not necessarily directed towards patent generation, but towards other kinds of research and progression, or even towards new R&D facilities and equipment.

The other two control variables that remain significant are those of general direct government expenditure towards R&D and the total R&D personnel per 1000 of the labor force. An increase in both on average also slightly increases the number of patent applications. This is expected, as direct expenditure towards R&D would help to fund R&D activities or provide better working conditions and facilities to R&D personnel. Direct funding would likely speed up the process of certain projects, resulting in a higher likelihood of patenting outcomes. Similarly, a higher number of R&D personnel in a country's labor force may lead to more knowledge sharing and generation of ideas. Although a higher quantity of R&D personnel may not necessarily be synonymous with quality of R&D personnel, it is more probable to yield a higher number of ideas and allow many minds to attempt to work on a certain project. Hence, this coefficient is higher than the other significant coefficient of general direct expenditure towards R&D, supporting the earlier discussed reasoning.

Additionally, the data was also tested for autocorrelation, as certain implications may be present for multiple regression models in cases of autocorrelation. The test for autocorrelation yielded a p value of 0.00, which is lower than the alpha of 0.05. Hence, the Null Hypothesis of there being no presence of autocorrelation in the data can be rejected, and the research can claim presence of autocorrelation in the chosen data set. This may indicate that predictions made by the regressions are not fully accurate and the estimates are not precise. Moreover, autocorrelation can result in underestimated standard errors, leading to misleading significance results, making this a valuable consideration for the limitations of the results. Nevertheless, such outcomes were anticipated, considering that it is plausible that the results in some years in the dataset are correlated with other years. Considering the general trend of a population increase in most countries in the years 2000-2020, it is expected that as the labor force expands, more people would participate in R&D, hence increasing the total R&D involvement. It is also unlikely that business enterprises would largely increase or decrease their R&D teams throughout the years, unless of instances like major layoffs. Thus, a business enterprise would retain a similar number of staff in a year compared to the previous, ultimately leading to the diversity of an R&D team in a current year being correlated to diversity in that R&D team in the previous year(s).

Sub question 1

To answer the first sub question, different patent categories were independently evaluated in relation to the BIH level (see Appendices E and F). Little can be concluded from the regression results, considering that none of the BIH coefficients are significant at either of the three significance levels. In 3 out of the 8 patent categories, there is on average an increase in the number of patent applications with higher diversity. These categories are those of fixed construction, mechanical engineering and electricity. Interestingly, these fields are the ones with the lowest ratio of male to female workers, with engineering in particular being having the least female workers and graduates out of all other fields (Neurosci, 2019). Naturally, this would imply these fields have lower levels of gender diversity. Meanwhile, the rest of patent categories yield negative association results, as also observed in Table 1, with the coefficients being much smaller and not significant unlike those in Table 1.

Sub question 2

In accordance with previously conducted literature on the topic, the research also aims to test whether there is an optimal level of BIH which would lead to the number of patent applications. When splitting BIH levels into categories, it can be observed in Table 2 that BIH 2, hence being 0.300-0.350, on average yields the highest number of patent applications (see Appendix G for full table). The comparison is done in relation to the lowest level of BIH in the dataset, ranging from 0 to 0.250. Thus, in relation to the lowest level of diversity, a BIH of 0.300-0.350 on average has the highest number of patent applications.

Moreover, the BIH coefficients decrease with every additional increase in the BIH category, hence yielding on average a progressively lower number of patent applications than the previous category, when comparing to the reference category of least gender diversity in R&D teams. Nevertheless, it is valuable to note that only the coefficient of BIH 3 is also significant at a 0.05 level, with the rest of BIH coefficients being insignificant despite the general trend. Thus, the results obtained from Table 2 and Appendix G corroborate the findings of Table 1 in which on average, higher gender diversity in R&D teams is not associated to a higher number of patent applications. Nevertheless, some level of gender diversity may be desirable to lead to the optimal increased number of patent applications. However, this level of diversity is rather low, considering that the most diverse population according to the BIH is one with BIH of 1. Such BIH would indicate that if randomly choosing an individual from a population, the chances the individual is female or male are equal.

Table 2

The individual effect of different BIH categories on the number of filed patent applications.

Variable	**Coefficient**
BIH 1	24.903
	[24.495]
BIH 2	61.379***
	[16.458]
BIH 3	32.631**
	[14.820]
BIH 4	21.196

	[17.671]
BIH 5	11.573
	[22.426]

Note: * indicates p < 0.1, ** indicates p < 0.05, *** indicates p < 0.01, standard errors are depicted in brackets. BIH categories follow as: BIH 1 0.250-0.300; BIH 2 0.300-0.350; BIH 3 0.350-0.400; BIH 4 0.400-0.450; BIH 5 0.450-0.500, with 0-0.250 category being the reference category. The table is a condensed version of the full table presented in Appendix G. It is missing the control variables, the constant, R squared and the number of observations, all of which can be found in the full version. The results are for OECD and non-OECD countries in the years 2000-2020, obtained through a time series regression, performed to answer Sub question 2.

To corroborate the results obtained in Table 2 regarding the optimal threshold for gender diversity in R&D teams, methodology from previously discussed studies was utilized. Thus, sub question 2 also aims to test for the non-linear nature of the relationship, as observed in the studies of Capozza and Divella (2023), Gonzalez-Moreno et.al. (2018), Neves Cavazotte and de Oliviera Paula (2021) and Liu, Ozer and Zhou (2023). Considering these studies not only propose nonlinearity, but also assume a U-shaped relationship between gender diversity and patenting, a quadratic regression has been applied (see Appendix H). To find the optimal level of gender diversity, the local maximum of the quadratic regression equation from Appendix H was calculated. To do this, a scatterplot of BIH levels and number of patent applications, was made, with the calculated quadratic regression (see Appendix I).

The local maximum of the regression was calculated to be 0.321, with this level of BIH on average yielding the highest number of patent applications, according to the quadratic regression. This level of BIH corroborates the results obtained with Table 2, where the BIH yielding the greatest number of patent applications was found to be in the range of 0.300-0.350, with 0.321 being in that range. This range of gender diversity is thus optimal for companies that wish to maximize their innovative output and the number of patent applications.

Nevertheless, when examining the Figure in Appendix I, two major outliers in the approximate range of BIH 0.300-0.350 can be observed, one at approximately 800 and the other at almost 1000 patent applications. Such outliers have the potential to influence the quadratic regression equation results and therefore the local maximum of the curve. It is therefore possible that the optimal range of BIH obtained is solely due to the presence of the two large outliers in the data. Nevertheless, the analysis will not take the outliers out of the dataset, as there is no reason to assume that these data points are untrue, and therefore are still valuable data points gathered from the dataset. Not representing the outliers would provide an unrealistic dataset, even in the case such dataset would yield different results and implications.

Conclusion and Discussion

Main Findings

Based on the performed analysis, business enterprises with a higher diversity index on average have a lower number of patent applications, as was concluded based on all models explored in the analysis. In terms of the chosen measure of the level of diversity in this research – the BIH, the results indicate that on average, as heterogeneity within R&D teams increases, hence becoming more gender diverse, the number of patent applications decreases. Therefore, regarding the central research question of "To what extent is higher gender diversity in research and development teams associated with increased innovative patenting within business enterprises?" it can be concluded that higher gender diversity in R&D teams is associated with increased innovative patenting within business enterprises only up to a certain threshold of diversity, and after, it is instead associated with a decrease in innovative patenting.

The most optimal level of gender diversity of such threshold, which would maximize innovative patenting is one in the range of BIH of 0.300-0.350, hence, a relatively low level of gender diversity. This means that if randomly selecting two members from an R&D team, the probability that they have a different gender is between 0.300 and 0.350, with women therefore representing less than 50% of the team. Such level is associated with the highest number of patent applications compared to the least gender diverse category of BIH. The number of patent applications on average decreases as the BIH deviates further from 0.300-0.350 and increases to 0.5, which is the highest level of BIH observed in the dataset. Hence, this further confirms that higher gender diversity in R&D teams is not associated with increase innovative patenting. Such conclusion is also corroborated by testing the U-shaped relationship between the two variables, which yields the most optimal level of BIH to be 0.321, being in the range of 0.300-0.350. Nevertheless, the presence of the U-shaped relationship cannot be validated, considering the presence of outliers in the dataset.

Additionally, higher gender diversity in R&D teams may be associated with increased innovative patenting in some patent categories. This is the case for the patent categories of fixed construction, mechanical engineering and electricity, but not the case for the other categories of human necessities, performing operations and transporting, chemistry and metallurgy, textiles and paper and physics. However, due to the insignificance of coefficients obtained through the chosen methodology, the association of the relationship between patent categories and gender diversity in

R&D teams cannot be confidently claimed, leaving these findings open to interpretation.

Discussion

The results of the analysis and the main findings are largely contrary to the existing literature on gender diversity which predominantly supports a positive association and between gender diversity in R&D teams and innovative activity of business enterprises. Particularly, this is contrasting to the findings of Martinez et al. (2017), who find both a positive and highly significant relationship between gender diversity in R&D teams and radical innovations. Similarly, they differ from Turner (2009), who concludes that by-products of gender diversity, such as better inter-personal communication skills and complex problem-solving abilities, all contribute to higher rates of innovative patenting.

The potential reasons of the contrasting findings regarding a negative association can be attributed to some limitations of the research. Nevertheless, certain literature does support the view that gender diversity in teams does not always lead to the most optimal outcomes. This is particularly true for identification of the optimal level of gender diversity, pointing to the BIH in the range of 0.300-0.350 to be the optimal level of patent maximization. Kou et.al (2019) not only also find a negative association between higher female personnel in Chinese research institutes and their R&D efficiency, but also provide psychological theoretical explanation for their findings. Diverse groups may be more prone to conflicts and tensions within the team, which in the long run can detrimentally impact team performance, particularly if remaining unsolved. Conflict has the potential to have negative impacts on idea sharing and collective work, supporting the findings of this paper's analysis.

Wikhamn and Wikhamn (2019) utilize the social categorization theory, the essence of which may be aligns with the results obtained via this research. By categorizing themselves into abstract teams based on gender divisions, social categorization may lead to involuntary and forced communication. Such dynamic is highly unfavorable within an R&D team, not only preventing voluntary exchange of ideas, but also promoting a non-sharing, closed off culture which decelerates innovative thinking. Thus, the findings of this research may reflect the difficulties gender diversity introduces into the workplace and particularly into R&D teams. Considering that most companies do not have a constantly changing team, introducing new team members may in fact already act as a deterrent and barrier to successful communication. It usually takes time for new team members to do accustomed to a new team, so potentially it is expected that introduction of new team members will at least in the beginning lead to stagnant communication and work ethics, with this being independent of the gender of the team members. Such conclusions are drawn in other studies, including that of Diaz-Garcia et al. (2013), who attribute the potential disadvantages of gender diversity to a need for greater coordination, interaction and communication within teams.

Additionally, studies focusing on the curvi-linear relationship between gender diversity and innovative performance also aligns with the findings of the paper's analysis. Despite the potentially flawed representation in of the relationship due to the presence of outliers, the analysis does conclude that a certain level of gender diversity within R&D teams maximizes patent applications, even though this level is relatively low, with BIH in the range of 0.300-0.350. Studies by Liu et. al. (2023); Gonzalez-Moreno et. al. (2018) and Neves Cavazotte and de Oliviera Paula (2020) all arrive to results which point to optimal levels of diversity to be lower than those of an approximate half to half split of a female and male personnel. All authors also study the relationship between gender diversity in R&D teams and innovative outcomes, further supporting the results of this paper. Thus, the relatively low optimal threshold for patent applications obtained in this paper may be due to a loss of synergies associated with higher diversity. Additionally, the limitation of not including any other diversity variables in the research, such as age, sexuality, or nationality may be also captured by the diminishing benefits of gender diversity, as described further in the Limitations section of this paper.

The negative association between gender diversity in R&D teams and innovative patenting applications concluded in this research also needs to be taken with consideration of the significance of contextual factors. Cultural and regional differences likely influence the extent to which gender diversity can impact a team, along with communication and work ethic. The data of this paper comes from varying countries, being both OECD and non-OECD and ranging over multiple continents. Hence, many cultural customs and company culture variables are difficult to evaluate and collect data on. As also emphasized by Na and Shin (2019), innovativeness of teams greatly relies on the country's initial perceptions and attitudes towards working and collaborating. It is therefore difficult to attribute gender diversity to be solely beneficial or detrimental to the patenting outcomes of R&D teams.

Limitations of the Research

The findings of the research may be due to certain shortcomings and limitations of both the utilized methodology and the dataset. The data is taken on a macro-level, which may not be the most optimal for the nature of the study. There may be large within-country heterogeneities, making micro-level data more appropriate by not having to account for such heterogeneities. However, due to the availability of data, a macro-level study was carried out, limiting the power of predictions made by the results. This brings about the issue of omitted variables bias, as it is impossible to fully control for all variables that impact both patent applications in a country and gender diversity in R&D teams, and in fact, these variables and their role may vary between countries. Despite the analysis suggesting a primarily negative association between gender diversity in R&D teams and innovative patenting, it is unfounded to conclude that gender diversity is the sole contributor to patent applications. The Data section mentions certain control variables the availability of which would greatly benefit the research, yet remains unavailable to open access and utilization. Not having controlled for these variables, its association cannot be confidently claimed, as the magnitude of the impact they have on the dependent and independent variables remains uncertain.

In particular, the issue of reverse causality may provide some explanation for the obtained results. Innovative firms whose company policy centers around innovative output may be more likely to attract male staff, possibly due to their prestige or higher paid salaries. Potentially, it is the firm's internal policies and company culture which attracts specifically male personnel, and such firms are on average more innovative than others. Since many innovative firms are found in high-tech industries, they may develop a company culture which aligns more with male-dominated fields or activities more favored by the male personnel. Additionally, highly successful innovative firms, like other business enterprises, tend to rely on networking for recruitment of new personnel. Considering the networks are predominantly male, they would likely appeal more to men, furthering the disproportionate gender balance in recruitment. This further reinforces the probability of the presence of reverse causality between the main examined variables in the research.

Moreover, considering that the presence of autocorrelation was confirmed through hypothesis testing in the panel data, this may also have implications for the findings. It is difficult to determine true association between gender diversity and innovative patenting in the case that gender diversity and patenting is correlated with the previous periods. For instance, if a country has low numbers of patenting applications and a low level of diversity, it is likely that both variables will remain approximately similar in the following year. Firms typically have a fixed budget and a fixed personnel count that does not significantly fluctuate yearly, unless of major disruptions such as layoffs. Thus, majority of business enterprises would retain a largely equal R&D team year by year, with the team not growing by large amounts. Unless the team happens to grow, there is little ability to introduce more gender diversity, other than by replacing current male members with female members. Even though the number of patenting applications may grow due to increases in the state of technology or new equipment and R&D facilities, gender diversity within R&D teams may remain relatively stable. Similar can be regarded in terms of patent applications, which also would unlikely largely increase yearly unless of knowledge breakthroughs, improved R&D facilities or existence of new technologies. Such presence of autocorrelation in the dataset can therefore bias results to appear like lower gender diversity is associated with higher number of patent applications, when in reality patent applications would increase in countries regardless of any changes in gender diversity within R&D teams.

In fact, it is of use to critically examine one of the assumptions made in this research - that diversity outcomes are attributed to gender diversity, rather than other kinds of diversity. It is possible that racial or geographical diversity are factors associated with patents and innovation but are instead captured by gender diversity – in the case that female workers would also provide other types of diversity, or countries with more female workers also have more diverse teams with regards to other types of diversity. The models utilized in the analysis do not control for any such diversity factors, due to a lack of data availability on them. While this connects to the issue of omitted variable bias, it is somewhat of a differing limitation, as gender diversity may already capture some of these other diversity factors. However, this research only aims to focus on gender diversity in R&D teams and its association with innovative patenting, ignoring potentially significant effects of other kinds of diversity within innovation teams and personnel of business enterprises.

Another valuable insight is the potential presence of a lag between gender diversity in R&D teams and innovation outcomes. It may take time for benefits of gender diversity to take full course and be pronounced. Considering that one of the main advantages of gender diversity is broader idea sharing and communication, such exchanges might not happen immediately, as the team gets to know new members and the members get accustomed to the company. Although the dataset utilizes panel data, this does not mean it fully captures the potential lags in the association between gender diversity and innovation outcomes.

Lastly, the research is limited by measuring only one innovation output - the number of patent applications. While the motivation behind using this variable and its advantages has been outlined in the Data section, it is by far not the only way to quantify innovation outcomes. Some innovation may not be captured by patents, considering that not all innovations are patented in the first place. Moreover, different countries may have varying laws regarding patent applications, making the applications process more cumbersome and less worthy for some countries. The quantity of patents also may not necessarily indicate the extent of innovation, as high impact or disruptive technology patents arguably demonstrate innovation better than simply the quantity of patents. Innovation can be captured in numerous other ways, such as new product launches, the changes in performance measures of a business enterprise, or customer satisfaction rankings. Such metrics also showcase innovation, although may be less direct indicators than patent applications. This insight is valuable to consider when evaluating not only the limitations of the research, but also when outlining potential ideas for further research.

Implications and Further Research

The findings of the research and varying perspectives on the influence of gender diversity in R&D teams on performance make it difficult to provide concrete recommendations to business enterprises. It is of use to consider that recently, companies and business enterprises have been becoming more diverse, including the introduction of gender diversity. Particularly, many larger and innovative companies now continuously release open statistics and reports on diversity in their workforce, oftentimes labeling themselves as pioneering for fostering a diverse team. It would therefore be unwise to claim that companies should aim for some level of diversity within R&D teams, as this has already been the trend for many business enterprises in the last decade. Instead, it is wise for companies to avoid extremes in their level of diversity, whether is it high or low extremes. Having highly homogeneous teams would likely prevent members from generating innovative ideas and patents due to a similar way of thinking and generally similar experiences in life and upbringing, all of which can potentially shape one's thought process. Meanwhile, the potential danger of having overly heterogenous teams lies in a high probability of conflict and discourses associated with extreme differences among team members, as earlier described in psychological evidence.

Moreover, regardless of the nature of the relationship between gender diversity and R&D, it prompts the need for effective diversity management within business enterprises. Considering that diversity may lead to increased coordination issues and conflict, yet has potential to enhance companies' capabilities, companies should dedicate resources to team bonding and conflict resolution strategies. This may include introducing dedicated personnel to deal with diversity and conflict related challenges, or improving management practices to better facilitate the downsides associated with different types of diversity. Such strategy can also entail providing diversity training and fostering an inclusive culture in business enterprises. An inclusive culture minimizes the risk of categorization based on similar group attributes and encourages less bias toward team members, ultimately encouraging mutual collaboration, respect, and idea sharing. All these qualities are positive towards a working environment, having potential to therefore maximize innovative outputs such as patents.

Taking this into consideration, conducting a study which compares business enterprises in regards to their conflict, diversity, and team building policies and their innovative output can be of interest when reflecting further on the research question examined in this paper. Since companies have differing policies on conflict resolution and diversity quotas, it may be valuable to analyze the effect of these variables on innovative outcomes like patents. It would be beneficial to conduct a causal study, rather than one of explanatory nature which aims to capture association. A causal study may include a treatment and control group, with the treatment group being a firm that introduces various conflict, diversity and team building policies and the control group not having any. The firms would have similar resources and be in the same product category, both introducing more gender diversity into their R&D teams. Various innovation outcomes can then be monitored in both firms over a certain time period, to provide evidence for any of the policies on R&D outcomes with increased gender diversity. Such study can provide strategic implications, particularly for company policy and potentially the human resource department, by offering ways in which gender diversity will yield the most optimal outcomes for business enterprises.

Additionally, the results only further emphasize the need for companies to encourage and support female participation in R&D roles. Such a strategy can help to break down potential bias against women in these positions and make female workforce participation more normalized, particularly in countries whose work cultures are tailored towards male personnel. This can be achieved through active recruitment programs, career development opportunities or special events dedicated towards female empowerment and participation in the field. It is of use for business enterprises to continuously identify potential barriers which prevent women from advancing in their profession. Barriers in the business environment which prevent successful female participation ultimately prevent optimal outcomes of knowledge sharing and generation of new ideas. Firms should therefore progress towards removing these barriers to not only increase female participation, but ultimately productivity and innovation by facilitating talent expansion.

These implications may be more difficult to implement, depending on the country the business enterprise is located in. Considering the role of social and cultural factors, certain places may have more enforced biases and barriers for women in the workplace. Hence, these would have a greater impact on the teams' willingness to successfully incorporate gender diversity and its benefits for the company's success. Thus, studying individual perceptions towards gender diversity in R&D teams between different countries may help to pinpoint whether it is the cultural and social persecutions which prevent gender diversity from benefitting business enterprises to the full extent. Performing a comparative cross-country analysis would indicate whether countries that generate most innovations are characterized by a certain work culture and specific attitudes towards different team members. Such study would also open the potential towards offering more tailored solutions per country of a business enterprise, depending on its local content. These findings would be especially valuable as developing countries continue to establish more business enterprises and increase their female participation in research roles, while developed countries look for more unique ways to generate more innovative output.

References

Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and innovation count as measures of regional production of new knowledge. *Research Policy*, *31*(7), 1069–1085.

Belova, G., & Ivanova, A. (2022). EU institutions: Revisiting gender balance and women's empowerment. *Laws*, *12*(1), 3.

Brennan, T., Ernst, P., Katz, J., & Roth, E. (2020, November 3). *Building an R&D strategy for Modern Times*. McKinsey & Company.

Capozza, C., & Divella, M. (2023). Gender diversity in European firms and the R&D-Innovation-Productivity Nexus. *The Journal of Technology Transfer*.

Charlesworth, T. E. S., & Banaji, M. R. (2019). Gender in science, technology, engineering, and mathematics: Issues, causes, solutions. *The Journal of Neuroscience*, *39*(37), 7228–7243.

Chen, G., Farh, J.-L., Campbell-Bush, E. M., Wu, Z., & Wu, X. (2013). Teams as innovative systems: Multilevel motivational antecedents of innovation in R&D teams. *Journal of Applied Psychology*, *98*(6), 1018–1027.

Czarnitzki, D., & Hottenrott, H. (2009). R&D investment and financing constraints of small and medium-sized firms. *Small Business Economics*, *36*(1), 65–83.

Díaz-García, C., González-Moreno, A., & Jose Sáez-Martínez, F. (2013). Gender diversity within R&D teams: Its impact on radicalness of Innovation. *Innovation Management: Policy and Practice*, *15*(2), 149–160.

European Institute for Gender Equality (2022), European Union Objectives for Gender Equality in Research and Innovation, European Commission.

Garcia Martinez, M., Zouaghi, F., & Garcia Marco, T. (2016). Diversity is strategy: The effect of R&D team diversity on innovative performance. *R&D Management*, *47*(2), 311–329.

Gruszczyński, M. (2020b). Women on boards and firm performance: A microeconometric search for a connection. *Journal of Risk and Financial Management*, *13*(9), 218.

Guellec, D., & Van Pottelsberghe De La Potterie, B. (2003). The impact of public R&D expenditure on business R&D*. *Economics of Innovation and New Technology*, *12*(3), 225–243.

Hall, B. H., Lotti, F., & Mairesse, J. (2009). Innovation and productivity in SMEs: Empirical evidence for Italy. *Small Business Economics*, *33*(1), 13–33.

Hall, B. H., Mairesse, J., & Mohnen, P. (2010). Measuring the returns to R&D. *Handbook of the Economics of Innovation, Volume 2*, 1033–1082.

He, X., & Jiang, S. (2019). Does gender diversity matter for green innovation? *Business Strategy and the Environment*, *28*(7), 1341–1356.

Ionescu, A., & Dumitru, N. R. (2015). The role of innovation in creating the company's competitive advantage. *Ecoforum Journal*, *4*(1).

Jones, G., Chirino Chace, B., & Wright, J. (2020). Cultural diversity drives innovation. Empowering Teams for Success. *International Journal of Innovation Science*, *12*(3), 323–343.

Katila, R. (2000). Using patent data to measure innovation performance. *International Journal of Business Performance Management*, *2*(1/2/3), 180.

Kou, M., Zhang, Y., Zhang, Y., Chen, K., Guan, J., & Xia, S. (2019). Does gender structure influence R&D efficiency? A regional perspective. *Scientometrics*, *122*(1), 477–501.

Liu, Z., Ozer, M., & Zhou, K. (2023). The role of status diversity in the innovative performance of R&D teams. *R&D Management*, *54*(1), 60–75.

Mazza, T., Furlotti, K., Medioli, A., & Tibiletti, V. (2023). Mandatory gender quotas: The impact on board and Committee meetings. *Management Research Review*, *47*(1), 45–63.

Moreno, Á. G., García, C. D., & Martínez, F. J. (2018). R&D team composition and product innovation: Gender diversity makes A difference. *European Journal of International Management*, *12*(4), 423.

Mowery, D. C., & Rosenberg, N. (2000). *Paths of innovation: Technological change in 20th-century America*. Cambridge University Press.

Na, K., & Shin, K. (2019). The gender effect on a firm's innovative activities in the emerging economies. *Economic and Business Aspects of Sustainability*, *11*(7), 1992.

Nadeem, M., Bahadar, S., Gull, A. A., & Iqbal, U. (2020). Are women eco-friendly? board gender diversity and environmental innovation. *Business Strategy and the Environment*, *29*(8), 3146–3161.

Nagaoka, S., Motohashi, K., & Goto, A. (2010). Patent statistics as an innovation indicator. *Handbook of the Economics of Innovation, Volume 2*, 1083–1127.

Nandal, N., Kataria, A., Dhingra, M. (2020). Measuring Innovation: Challenges and Best Practices. *International Journal of Advanced Science and Technology, 29*(5), 1275-1285.

OECD (2023), Joining Forces for Gender Equality: What is Holding us Back?, OECD Publishing, Paris.

Organisation for Economic Co-operation and Development. (n.d.). *OECD.Stat*. OECD Statistics.

Paulus, P. B., van der Zee, K. I., & Kenworthy, J. B. (2019). Diversity and group creativity. *The Oxford handbook of group creativity and innovation*, 33-49.

Pfeifer, C., & Wagner, J. (2013). Is innovative firm behavior correlated with age and gender composition of the workforce? evidence from a new type of data for German Enterprises. *Journal for Labour Market Research*, *47*(3), 223–231.

Rogers, M. (1998). *The definition and measurement of innovation.* (Melbourne Institute Working Paper; No.10/98). Melbourne Institute of Applied and Social Research.

Teruel, M., & Segarra-Blasco, A. (2017). The Link Between Gender Diversity and Innovation: What is the Role of Firm Size? *International Review of Entrepreneurship, 15*(3), 319-340.

Turner, L. (2009). Gender diversity and innovative performance. *International Journal of Innovation and Sustainable Development*, *4*(2/3), 123.

Wikhamn, W., & Wikhamn, B. R. (2019). Gender Diversity and Innovation Performance: Evidence from R&D workforce in Sweden. *International Journal of Innovation Management*, *24*(07), 2050061.

Appendices

Appendix A

Descriptive statistics of all variables.

Variable	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**	**Observations**
# of patent applications	217.466	212.652	4.8	958.7	364
BIH	0.378	0.067	0.215	0.5	380
Gross domestic expenditure on R&D	4872.027	6654.287	17.776	39396.79	424
Government budget allocation for R&D	0.465	0.194	0.11	1.04	408
Business enterprise researchers as % of national total	37.492	16.187	6.2	72.26	430
Total R&D personnel per 1000 of employment	9.035	4.769	1.43	24.31	431
R&D tax expenditure and direct government funding	0.042	0.058	0	0.284	367

Note: The table includes descriptive statistics of all the variables that were used to answer the main research question.

Appendix B

Ranking of countries.

Variable	**Country**	
	Minimum	**Maximum**
# of patent applications	Latvia	Taiwan
	Estonia	Italy
	Slovakia	Belgium
BIH	Taiwan	Latvia
	Czechia	Russia
	Hungary	Romania
Gross domestic expenditure on R&D	Latvia	Taiwan
	Estonia	Russia

	Slovakia	Singapore
Government budget allocation for R&D	Chile	Singapore
	Romania	Taiwan
	South Africa	Russia
Business enterprise researchers as % of national total	Argentina	Taiwan
	Latvia	Ireland
	South Africa	Russia
Total R&D personnel per 1000 of employment	Chile	Taiwan
	South Africa	Belgium
	Argentina	Norway
R&D tax expenditure and direct government funding	Estonia	Hungary
	Latvia	Ireland
	Chile	Portugal

Note: The table includes the ranking of all 21 countries used in this research. The average for all variables used to answer the main research question has been calculated. For each of the variables, the table ranks countries based on the 3 countries with the lowest averages per variable (minimum) and 3 countries with the highest averages per variable (maximum).

Appendix C

Hausmann Test of fixed and random effect models.

Variable	Coefficients		Difference
	Fixed	Random	
BIH	-337.236	-373.723	21.318
Chi Squared		0.089	

Note: The Hausmann test represents coefficients of the impact of BIH onto the number of filed patent applications in both OECD and non-OECD countries in the years 2000-2020. The difference is calculated by subtracting the BIH random effects coefficient from the BIH fixed effects coefficient.

Appendix D

Time series regressions of control variables.

Variable	Control Variable: General direct expenditure	Government budget	% of total labor force	Total per 1000 of labor force	Tax and funding
BIH	-282.226**	-375.160***	-318.199**	-311.903***	-359.063***
	[111.260]	[125.704]	[126.470]	[120.511]	[123.813]
General direct expenditure	0.024***	0.023***	0.022***	0.022***	0.022***
	[0.001]	[0.002]	[0.002]	[0.002]	[0.003]
Government budget		17.646	13.748	-93.796*	-2.835
		[52.413]	[51.946]	[53.218]	[56.059]
% of total labor force			1.166**	-0.153	0.291
			[0.474]	[0.511]	[0.626]
Total per 1000 of labor force				12.590***	10.986***
				[2.299]	[3.269]
Tax and funding					-91.371
					[153.031]
Constant	310.675***	339.437***	282.963***	272.413***	207.756***
	[64.324]	[71.766]	[75.502]	[73.231]	[72.672]
R Squared	0.409	0.452	0.492	0.480	0.571
Observations	312	302	302	302	249

Note: * indicates p < 0.1, ** indicates p < 0.05, *** indicates p < 0.01, standard errors are depicted in brackets. Control variables are added each one by one individually into a time series regression, representing the impact of BIH onto the number of patent applications in OECD and non-OECD countries in the years 2000-2020.

Appendix E

The individual effect of different patent categories on the number of filed patent applications.

	Patent Category			
Variable	A	B	C	D
BIH	-25.853	-5.201	-0.603	-1.301
	[59.776]	[57.378]	[33.647]	[19.063]
General direct expenditure	0.013***	0.009***	0.001**	0.001***
	[0.001]	[0.001]	[0.001]	[0.000]
Government budget	90.128***	-5.258	51.516***	21.318***
	[25.464]	[26.413]	[14.333]	[8.197]
% of total labor force	1.205***	0.290	0.534***	0.228**
	[0.291]	[0.299]	[0.164]	[0.093]
Total per 1000 of labor force	-7.786***	-1.190	-2.468***	-0.558
	[1.524]	[1.637]	[0.868]	[0.481]
Tax and funding	330.031***	127.591	85.043**	-39.179**
	[63.945]	[65.622]	[36.005]	[19.960]
Constant	36.890	55.456	43.322	-5.659
	[49.203]	[52.680]	[29.249]	[10.702]
R Squared	0.407	0.307	0.157	0.182
Observations	299	281	299	292

Note: * indicates p < 0.1, ** indicates p < 0.05, *** indicates p < 0.01, standard errors are depicted in brackets. Patent categories follow as: A – human necessities; B – performing operations and transporting; C – chemistry and metallurgy; D – textiles and paper. The effect is obtained by running a time series regression per each individual patent category for OECD and non-OECD countries in the years 2000-2020, and is performed to answer Sub question 1.

Appendix F

The individual effect of different patent categories on the number of filed patent applications.

	Patent Category			
Variable	E	F	G	H
BIH	28.026	10.501	-1.343	7.495
	[20.771]	[38.746]	[40.277]	[36.178]
General direct expenditure	0.002***	0.006**	0.008***	0.004***
	[0.000]	[0.000]	[0.000]	[0.001]
Government budget	0.013***	0.009***	0.001**	0.001***
	[0.001]	[0.001]	[0.001]	[0.000]
% of total labor force	0.122	0.314	0.105*	0.491*
	[0.100]	[0.188]	[0.195]	[0.173]
Total per 1000 of labor force	-0.313	-1.293	-0.748	-1.466
	[0.529]	[0.986]	[1.024]	[0.910]
Tax and funding	11.103	34.900	111.946**	57.324
	[22.178]	[41.368]	[42.946]	[38.280]
Constant	-6.189	0.538	20.684	5.764
	[14.099]	[26.195]	[25.343]	[23.392]
R Squared	0.369	0.409	0.579	0.414
Observations	286	270	290	297

Note: * indicates p < 0.1, ** indicates p < 0.05, *** indicates p < 0.01, standard errors are depicted in brackets. Patent categories follow as: E – fixed constructions; F - mechanical engineering, lighting, heating, weapons, blasting; G – physics; H – electricity. The effect is obtained by running a time series regression per each individual patent category for OECD and non-OECD countries in the years 2000-2020, and is performed to answer Sub question 1.

Appendix G

The individual effect of different BIH categories on the number of filed patent applications.

Variable	Coefficient
BIH 1	24.903
	[24.495]
BIH 2	61.379***
	[16.458]
BIH 3	32.631**
	[14.820]
BIH 4	21.196
	[17.671]
BIH 5	11.573
	[22.426]
General direct expenditure	0.023***
	[0.003]
Government budget	-85.266
	[55.756]
% of total labor force	0.199
	[0.626]
Total per 1000 of labor force	15.211***
	[3.400]
Tax and funding	-151.502
	[154.430]
Constant	77.751
	[56.608]

R Squared	0.518
Observations	277

Note: * indicates p < 0.1, ** indicates p < 0.05, *** indicates p < 0.01, standard errors are depicted in brackets. BIH categories follow as: BIH 1 0.250-0.300; BIH 2 0.300-0.350; BIH 3 0.350-0.400; BIH 4 0.400-0.450; BIH 5 0.450-0.500, with 0-0.250 category being the reference category. The effect is obtained by running a time series regression for BIH categories, for OECD and non-OECD countries in the years 2000-2020, and is performed to answer Sub question 2.

Appendix H

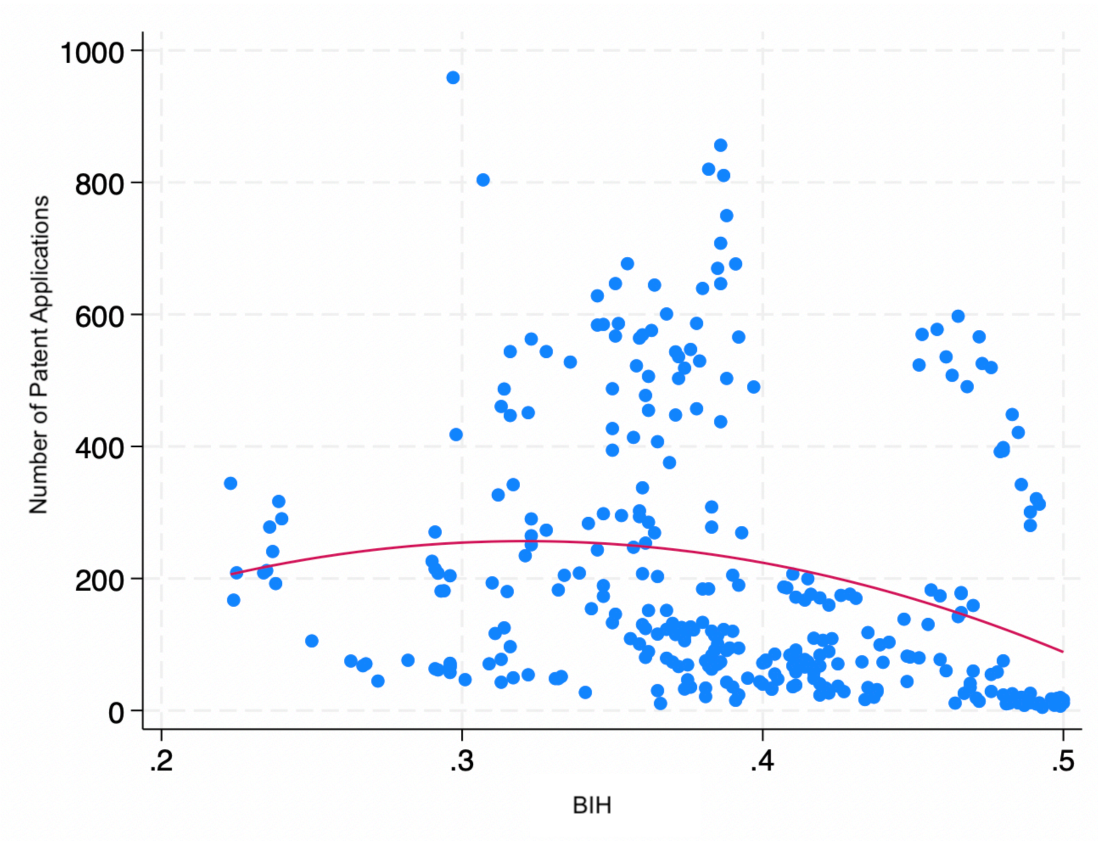
Quadratic model of the effect of different BIH categories on the number of filed patent applications.

Variable	Coefficient
BIH	3365.304**
	[1663.278]
BIH * BIH	-5242.637**
	[2177.565]
Constant	-283.4647
	[312.812]
R Squared	0.055
Observations	315

Note: * indicates p < 0.1, ** indicates p < 0.05, *** indicates p < 0.01, standard errors are depicted in brackets. The quadratic regression was run for OECD and non-OECD countries in the years 2000-2020, and is performed to answer Sub question 2.

Appendix I

The relationship between BIH and the number of filed patent applications.



Note: The relationship is presented by plotting raw data and running a quadratic regression from Appendix H. The relationship is represented for OECD and non-OECD countries for the years 2000-2020.