Norwegian School of Economics

Bergen, Fall 2020

The Patenting Gender Gap in Europe

Geographical patterns and educational gender gaps

Antoine François Martin

Supervisor: Associate Professor Steffen Juranek

Master thesis, MSc in Economics and Business Administration

NORWEGIAN SCHOOL OF ECONOMICS

This thesis was written as a part of the Master of Science in Economics and Business Administration at NHH. Please note that neither the institution nor the examiners are responsible − through the approval of this thesis − for the theories and methods used, or results and conclusions drawn in this work.

Preface

This master thesis was written as part of the Master of Science in Economics and Business Administration, major in International Business at the Norwegian School of Economics. During the completion time of this thesis, I have gained valuable knowledge and experience in the current gender gap existing in innovation and patenting in the European continent, as well as insights into the geographical patterns and the educational causes of this gender gap.

I would like to thank my supervisor Steffen Juranek, Associate Professor at the Norwegian School of Economics, for his excellent guidance and help throughout this project. Despite the unprecedented circumstances and difficulties caused by the COVID-19 pandemic, he was always available for providing assistance whenever I needed.

Abstract

With a focus on Europe, the present thesis empirically investigates the geographical patterns of the patent gender gap as well as the linkage between the patent gender gap and the education gender gap potentially explaining why women patent less than men. This paper contains a twofold analysis. Firstly, by a thorough spatial analysis of data covering 267 European regions, it is shown that, following a pattern similar to global patenting and innovation activity, women's patenting is more frequent in cities and their vicinities. Secondly, by developing beta regression and panel regression models, it is shown that a set of various education-related gender gaps interdependently affect the severity of the patent gender gap. Being of interest to innovation agencies and other innovation-policy organisations, the contribution of this study lies in the analysis of the geographical patterns of women's patenting as well as the intricate nature of the causes of a wide, although narrowing, gender gap in patenting and innovation.

Contents

Table of Contents

1. Introduction

After having had totally different and quite pejorative interpretations and meanings for centuries, Joseph Schumpeter made innovation win its spurs within economic theory in the mid 20th century. From that moment onwards, economic theoreticians began to see product and technological innovation as an engine of competitive advantages, a catalyst of economic growth and an essential driver of human progresses (Godin, 2019). Defined by the European Central Bank as the development and application of ideas and technologies that improve goods and services or make their production more efficient (2017), innovative processes and outbreaks benefit both consumers and businesses. As a driver of productivity improvements, innovation greatly contributes to economic growth in capitalist economic models: “creative destruction is the essential fact about capitalism” (Schumpeter, 1942, p.83), creative destruction being used as a term describing the disruptive transformation processes characterising innovation.

As a tool used by businesses, research organisations and individuals to protect and enforce their intellectual property rights, patents play an increasingly important role in innovation and economic performance, especially in the fields of information and communication technologies – ICT – and biotechnology (OECD, 2004). Patents play an important role throughout the entire technology lifecycle from Research and Development to market diffusion and allow their originators to derive financial gains after the licensing of their competitive technologies to third-party entities. More precisely, a patent legally grants its holder the right to prevent others from commercially exploiting their invention for a limited period of time[1] without the holder’s authorisation, typically granted through licensing. This protection applies in the geographical area where the patent application is filed[2]. Effectively, a patent turns an inventor’s know-how into a tradeable asset, enabling business growth opportunities and job creations (WIPO, n.d.).

[1] Maximum term of 20 years in Europe and the United States of America.

[2] For instance, a patent application filed at the European Patent Office will protect an invention across the European continent.

Despite an increase in the share of women in international patent applications filed via the World Intellectual Property Organisation between 1995 and 2015[3], it is empirically observed that a gender gap exists in patenting. Being a human right clearly stated by the United Nations as one of the seventeen Sustainable Development Goals[4], higher gender equality is commonly seen as a driver for both higher economic growth (Bertay, Dordevic, & Sever, 2020) and research productivity (Gui-Diby, Pasali, & Rodriguez-Wong, 2017). Moreover, diverse and inclusive teams are more innovative (Phillips, 2014) and diverse companies are economically more profitable (Hunt et. al, 2020). Therefore, the existence of a patent gender gap is likely to impair economic growth by missing out on the great ideas of the female population and on an important part of human productivity (Liberda & Zajkowska, 2017).

Moreover, gender inequalities are a long-lasting characteristic of a vast majority of societies, both today and in history. Among other elements, a crucial dimension through which gender inequalities manifest themselves is the gain of human capital through education, which in turn determines gender inequalities in a broad set of macroeconomic and social indicators, like employment, public life and the role in society of men and women. However, the education gender gap has been greatly closed – or even reversed – since the nineteenth century (Bertocchi & Bozzano, 2019). Hence, one can reasonably assert that the gender disparities in human capital accumulation are now narrower than they used to be. Concurrently to this closing in education, women have made a massive entrance into the western corporate world throughout the last fifty years. However, the underrepresentation of women in the corporate world is persisting and carries negative effects in the form of productivity loss as well as a lessened innovation capacity, for instance through patenting among other innovation means. In addition, women are more at risk to face a glass ceiling than their men counterparts, preventing women workers from attaining manager positions, hampering women's career advancement to decision-making and leadership roles (Miller, 2019) and limiting women empowerment. This directly and negatively affects equality at the workplace, while an equal corporate culture is known to be a major driver of innovation, much more effective than salary incentives. Fostering businesses' innovative mindset and culture through the improvement of equality at the workplace could lead to immense progress in economic growth, the potential gain in global

[3] From 17% in 1995 to 29% in 2015. See: https://www.wipo.int/pressroom/en/articles/2016/article_0015.html. These figures cover 151 countries worldwide.

[4] See: https://sdgs.un.org/goals.

GDP being estimated at $8 trillion in the ten-year period 2019-2028 (Shook & Sweet, 2019). Hence, the potential gains of narrowing the patenting gender gap – as part of the fostering of a broader innovation culture – are massive and potentially reinforce the virtuous circle of innovation (Akinyemi, 2016).

Focused on the geographical patterns of patenting by women and the implications of the education gender gap on the patenting gender gap in Europe, this paper is structured as follows. The second section gives an account of previous literature that constitutes grounds for the statement of the four main hypotheses investigated in this paper. The third section presents the data collection process and the datasets used in this study, as well as subsidiary statistics that do not constitute the focal point of this study. The fourth section describes the econometric tools and models used to investigate the last three hypotheses formulated in the second section. The fifth section empirically analyses the four hypotheses stated in section 2, using both spatial visualisation of the data – for the first hypothesis – and the econometric tools presented in section 4 – for the three remaining hypotheses. Finally, the conclusions that can be drawn from this study are presented in the sixth and final section.

These sections present several sets of results. Firstly, large cities seem to play a role in stimulating patenting by women, as the patenting gender gap is narrower in their surroundings, although far from an equal gender split. Secondly, narrowing the education gender gap seems to result in a narrower patenting gender gap on condition that the science, technical, engineering and mathematical fields of study are taken as the focus when narrowing the education gender gap. Thirdly and lastly, encouraging gender equality unsurprisingly result in a larger share of women among the total patentee population.

2. Literature support and hypotheses statements

Based on existing literature and prior research, four hypotheses on contributing factors to more women patentees can be formulated: a higher women's patenting activity located around large cities, the gender gap in the educational attainment level, the dominant academic discipline of the degree obtained and the associated gender gap, as well as gender equality at large coupled with the role of women in society and labour participation. Each hypothesis is preliminarily contextualised with supportive literature.

2.1 H1. City effect

The importance of cities in the innovation process and economic growth has been abundantly documented, as early as 1890 by the English economist Alfred Marshall who originated the theory of knowledge spillovers. Augmented successively by Arrow (1962), Romer (1986) and Glaeser et. al (1992), the knowledge spillovers theory argues that competitive or non-competitive positive externalities arise in specialised industrial geographical clusters, both within and across industries. Due to a geographical proximity[5], these spillovers stimulate innovation (Jaffe, Trajtenberg, & Fogarty, 2000), and, therefore, can stimulate patenting. Hence, cities can be seen as focal points of innovation and entrepreneurship processes (Florida, Adler, & Mellander, 2016) that aggregate innovation and production, activities that are co-dependent with regard to skills, knowledge, products and/or markets (Pratt, 2008). Completing the argumentation highlighting the importance of cities for innovation, Bettencourt et. al (2007) argue that patent activity correlates positively with the presence of cities and their size.

Considering the crucial role of cities in generating and stimulating innovation processes embodied by, for instance, patenting, the first hypothesis formulated is that cities also have a stimulating effect on patenting by women. Put differently, the hypothesis is that the share of women among patent applicants is higher in regions that contain large cities and/or a large industrial hub. Due to both the lack of a common definition of what is a large city and the variability of city sizes across European countries, this hypothesis will be tested and analysed in section 5 based on spatial visualisations of the data, i.e. maps, rather than regressions.

---

[5] Hence, the term cluster.

2.2 H2. Educational gender gap

Following a Granger causality relationship, there is a strong causality between the educational attainment level of an individual and the innovative capacity of this individual (Makkonen & Inkinen, 2013). Therefore, a higher educational attainment level achieved by a woman should boost her innovative capacity, which is here considered under its patenting form, i.e. the capacity of a woman to file a patent application. Following this logic, a woman holding a bachelor's degree should have a higher patenting capacity than a woman holding a lower-level degree or no degree at all. Similarly, the patenting capacity should increase as the educational attainment level increases, a master's degree and a PhD degree being the second highest and the highest degrees possibly achieved, respectively.

Considering the positive correlation between the patenting capacity and the educational attainment level described hereinabove, there should be a positive correlation between the narrowness of the educational gender gap and the patent gender gap, i.e. the share of women among patent applicants. A particular attention is given to the PhD gender gap, as the possession of a "title" is one of the two most important characteristics venture capitalists search for among scientists when they extend invitations to participate in scientific boards and research projects, the other important characteristic being high productivity, women having either of these less frequently than men (Stephan & El-Ganainy, 2006).

2.3 H3. STEM education gender gap

This hypothesis is to be considered in parallel with the former hypothesis, both as an extension and a deepening of the second hypothesis. The acquisition of practical knowledge can be seen as the first prerequisite in forming innovative – and therefore patenting – capacity. Practical knowledge is then leveraged in combination with science knowledge such as mathematics, physics, chemistry and biology, and this combination is deployed and used in the search for improved and new products and processes (Scott, 2006). Scott also identifies the new-economy sectors – such as high-technology manufacturing and R&D, biotechnology and IT technology and services – as particularly prone to innovation. An example of the importance of new-economy sectors in an innovation and patenting context is the fact that the second and third most important sources of patenting in Germany were respectively electrical machinery and mechanical elements, between 2015 and 2017, according to the World Intellectual

Property Organisation. Therefore, receiving education in a STEM[6] field can be seen as a driver of patenting activity, regardless of gender.

The third hypothesis formulated is that not only the narrowing of the educational gender gap correlates positively with the share of women in patenting – hypothesis 2 – but also that receiving education in a STEM-related field is another driver of women's patenting activity. Hence, there should be a positive effect when reducing the STEM-education gender gap on the patent gender gap. Therefore, a positive correlation is expected between the share of women among STEM students and the share of women patentees.

2.4 H4. Gender inequality and the role of women

According to the European Commission, women are underrepresented in the European labour market[7]. A dimension of gender inequality taken in a job market context is the relative inaccessibility to the labour market or to some sectors of this market for women compared to men, and gender disparities are particularly blatant in STEM employment[8]. Moreover, labour-related gender inequality can also be present in the working time discrepancies between genders, namely that women more often work part-time compared to men (European Commission, 2013).

The fourth hypothesis formulated is that the countries or regions with a lower gender inequality present a higher share of women among their population of patent applicants. The measure used is the Gender Inequality Index calculated by the United Nations Development Programme. A negative correlation between the share of women patentees and this index is expected, since a lower index value is a sign of a diminished gender inequality and the other way around. Another rationale for this hypothesis comes from the WIPO statements that

[6] Science, Technology, Engineering and Mathematics.

[7] See: https://ec.europa.eu/info/policies/justice-and-fundamental-rights/gender-equality/women-labour-market-work-life-balance/womens-situation-labour-market_en#gender-segregation-in-the-labour-market

[8] For instance, a third of men in Europe is employed in science, technology and engineering, while only 7% of women are employed in the same sectors. In contrast, a third of women in Europe is employed in education, health and social work, compared to only 8% of men in this sector. See: https://ec.europa.eu/info/policies/justice-and-fundamental-rights/gender-equality/women-labour-market-work-life-balance/womens-situation-labour-market_en#gender-segregation-in-the-labour-market

"closing the gender gap would benefit everyone" and "anything that restricts innovation (…) means we are all less well-off"[9].

Table 1 hereunder summarises the four hypotheses, the variables involved as well as the expected sign of the associated estimate coefficient. Further details and explanation about the definitions of variables are available in section 3.3.

	Hypotheses	Variables	Expected sign
1	H1	fem_share/Maps visualisation	+
2	H2	fem_share/F_share_tereduc_PhD	+
3	H3	fem_share/F_Share_STEM	+
4	H4	fem_share/GII	-

Table 1. Hypotheses summary

[9] See: https://www.wipo.int/ip-outreach/en/ipday/2018/innovation_creativity_gender_gap.html

3. Data

3.1 Countries considered

A sample of 27 European countries was included in the study. These countries are, in the alphabetical order: Austria, Belgium, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, the Republic of Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland and the United Kingdom[10]. They are highlighted in blue in figure 1 hereunder:

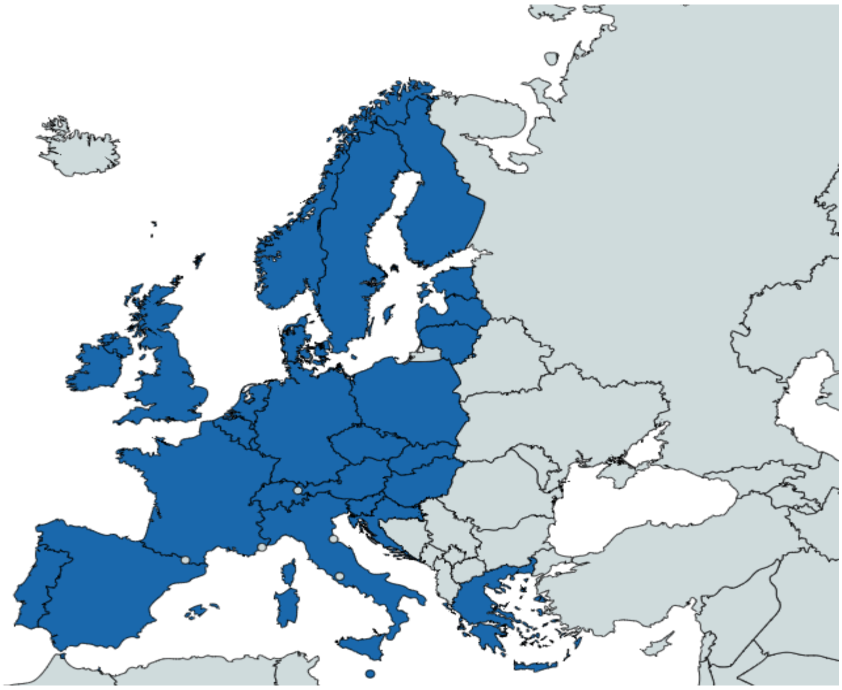


Figure 1. Countries included in the study

Several European countries were excluded from the sample. These countries were excluded from the study based on five reasons. The first reason for their exclusion is the missingness of data for these countries in the OECD and European Eurostat database, despite their membership to either the European Union or the European Economic Area. The second reason for exclusion is a country's non-membership to either the European Union or the European

[10] The United Kingdom was still a member of the European Union 31 January 2020. As of redaction date of this master thesis, the United Kingdom is still a member of the European Economic Area until 31 December 2020.

Economic Area. The third reason for exclusion is a debatable geographical location on the European continent. The fourth reason for exclusion consists in territorial conflicts and claims, splitting the country into two zones: an internationally recognised and undisputed national territory, and a disputed territory, possibly recognised by none, one or several other countries in the world. The fifth and final reason is a too small population size and land area, coupled with a close dependence to another country that is already included in the study. Hence, the following countries were excluded from the study based on the first of the above-mentioned reasons: Bulgaria, Cyprus, Iceland, Liechtenstein and Romania. The following countries were excluded based on the second reason: Russia, Belarus, Ukraine, Moldova, Bosnia Herzegovina, Albania, Montenegro, Serbia, Kosovo, and North Macedonia. Turkey was excluded from the study based on the third reason, while Cyprus was excluded based on the fourth reason. Finally, the microstates of Andorra, Liechtenstein, Monaco, San Marino and Vatican City were excluded based on the fifth and last reason.

3.2 OECD and WIPO databases, gender identification

The dependent variable considered throughout this study is the share of women patent applicants in European regions, in the sense of the European Union's Nomenclature of Territorial Units for Statistics[11] and at the NUTS2 level. To do so, patent data was first retrieved from the OECD's REGPAT database[12]. This database presents patent data linked to regions based on the address of the applicants and inventors. The patent data is originally regionalised at the NUTS3 level, the last and lowest level of the Nomenclature Territorial Units for Statistics (Maraut et al., 2008). The difference between NUTS2 and NUTS3 levels is simply the length of the string of characters of the regional name code. A NUTS2 level code comprises four characters – two letters followed by two digits[13] – while a NUTS3 level code comprises five characters – two letters followed by three digits – with the same first four characters for NUTS3 regions belonging to the same NUTS2 region. For example, the

[11] French: *Nomenclature des Unités Territoriales Statistiques*, hence abbreviated NUTS.

[12] Last version from July 2019.

[13] Except for Germany and the United Kingdom, where NUTS2 regions can be denoted by either a sequence of two letters and two digits or a sequence of three letters and one digit. For example: DEG0 or UKH1. This does not affect the way NUTS2 regional code is implied by the NUTS3 regional code.

Austrian NUTS3 regions AT111, AT112 and AT113[14] all belong to the Burgenland NUTS2 region with the code AT11. Therefore, the observations included in the REGPAT dataset were converted to the NUTS2 level by simply deleting the last digit of the NUTS3 code.

Considering that the gender of each patent applicant was a central point of focus for this study, the observations retrieved from the REGPAT database were matched with the patent data compiled in the OECD's HAN database[15], which provides a grouping of patent applicants' names. The REGPAT and HAN databases were consequently matched and merged by applicant id, so each patent application was assigned a name and a first name. However, the gender of each patent patentee was then still missing. To identify the gender of each patentee, I made use of the WIPO[16] worldwide gender-name dictionary which compiles 6.2 million names for 182 countries disambiguating the gender of patent applicants (Lax Martínez, Raffo, & Saito, 2016). The gender of each patent applicant was therefore identified, which enabled the calculation of the women share of patentees among all patentees in the NUTS 2 regions of the countries included in the study.

3.3 Datasets and variable definition

This paper goes deeper than the country level and focuses on the regional level, as there can be great discrepancies within each country. Put differently, the women share among patentees can vary greatly across regions within the same country. Hence, the last step in the construction of the dataset was to retrieve the women share among patentees at the regional level, more precisely at the NUTS2 level, based on the information extracted from the REGPAT database.

Two different datasets were constructed and are used in this paper: a cross-sectional data and a panel data, both covering 267 NUTS2 regions in the countries forming the sample. The cross-sectional dataset contains the women share among all patentees in 2017, while the panel dataset contains the women share among all patentees from 2013 to 2017, year by year within this time frame. Both datasets also include a set of ten variables retrieved from the Eurostat

[14] The Mittelburgenland, Nordburgenland and Südburgenland districts, respectively.

[15] Latest version from July 2019.

[16] World Intellectual Property Organisation.

database observed at the NUTS2 regional level, which relate to regional educational attainment and demographics. Five additional variables relating to both educational fields, innovation capacity and access to employment were included in the cross-sectional dataset. These additional variables were retrieved from the 2017 Global Innovation Index (GII) report, published by the WIPO in collaboration with Cornell University and INSEAD. Composed of 81 sub-indicators relating to political environment, education, infrastructure and business sophistication, the GII presents metrics about the innovation performance of 127 economies around the world. However, the data extracted from the 2017 GII were available at the country level only, not at the regional level. Hence, all NUTS2 regions of a same country were assigned the same value for the three variables retrieved from the 2017 GII report. Lastly, one additional variable was added: the Gender Inequality Index. Calculated by the United Nations as part of its Development Program, this index captures, among other areas, the gender inequality in access to the labour market and women empowerment. The Gender Inequality Index can therefore be used as a proxy for gender equality in the labour market and society. Data for the Gender Inequality Index was available at the country level only. Each NUTS2 region of a same country was thus assigned the same value for this variable. Overall, seventeen variables were retrieved and used in this study. Table 2 gives an overview of the seventeen variables considered, their respective definition and abbreviation.

	Variable	Variable Definition
1	fem_share	Women share among patentees, %
2	F_early_educ_count	Girls enrolled in early childhood education
3	F_share_early_educ	Share of girls enrolled in early childhood education, %
4	F_share_tereduc_Bach	Share of women in all students enrolled in a bachelor programme, %
5	F_share_tereduc_Mast	Share of women in all students enrolled in a master programme, %
6	F_share_tereduc_PhD	Share of women in all students enrolled in a PhD programme, %
7	F_share_pop	Share of women population, %
8	F_EDat02	Women aged 25-64 with less than primary and secondary education, %
9	F_EDat34	Women aged 25-64 with secondary and post-secondary education, %
10	F_EDat38	Women aged 25-64 with secondary, post-secondary and tertiary education, %
11	F_EDat58	Women aged 25-64 with tertiary education,%
12	Share_Grad_STEM	Share of STEM graduates among graduate population, %
13	F_Share_STEM	Share of women in STEM graduates, %
14	KnowlIntEmp	Share of knowledge-intensive employment, %
15	FemEmpdAdvDeg	Share of women employed with advanced degrees, %
16	GII	Gender Inequality Index
17	GInnIx	Global Innovation Index

Table 2. Variable definition[17]

[17] STEM = Science, Technology, Engineering and Mathematics.

Variable 1 – fem_share – measures the severity of the patent gender gap, variables 4, 5 and 6 measure the education gender gap at the bachelor, master and PhD level respectively, variable 13 measures the education gender gap within the STEM academic fields.

3.4 Subsidiary statistics: team sizes by gender

Importantly, the inventorship share[18] of each patentee within the same patent application is not considered in either dataset. A patent application can be filed by one applicant alone or several applicants jointly. In the latter case, there can be only one or several women, one or several women alongside one or several men, or only one or several men filing the same patent application. Put differently, this study does not account for the composition of each team applying for a patent, and a woman applying for a patent alone will be accounted for in the same way regardless of the other persons she may have filed the application with. However, some basic statistics about inventorship shares and team sizes can be retrieved.

Tables 3 and 4 show the percentage of team size for women and men patentees, respectively. Importantly, these tables do not account for the gender of co-patentees forming the rest of the team. Put differently, the row “one woman and two other” in table 3, i.e. two other patentees, covers teams composed of one woman and two men, teams composed of one woman plus another woman and a man and teams composed of three women. Similarly, the row “one man and two other” in table 4, i.e. two other patentees, covers teams composed of one man and two women, teams composed of one man plus another man and a woman and teams composed of three men.

[18] Defined as the share of a patentee within his/her patent applications. For example, a patentee who filed a patent application alone will have an inventorship share of 100% for this patent application, while a patentee who filed a patent application together with two other patentees will have an inventorship share of 33.33% for this patent application.

	Team Composition	Team Size	Share (%)
1	One Woman Alone	1	7.035
2	One Woman and One Other	2	16.843
3	One Woman and Two Other	3	13.775
4	One Woman and Three Other	4	17.829
5	One Woman and Four Other	5	12.479
6	One Woman and More Than Four Other	>5	32.039

Table 3. Team sizes for women patentees

	Team Composition	Team Size	Share (%)
1	One Man Alone	1	16.770
2	One Man and One Other	2	22.485
3	One Man and Two Other	3	7.921
4	One Man and Three Other	4	15.082
5	One Man and Four Other	5	8.983
6	One Man and More Than Four Other	>5	28.759

Table 4. Team sizes for men patentees

From these tables, it appears that women tend to file patent applications within larger teams than men. In particular, women filing patent applications alone are much less frequent than men filing patent applications alone, with 7.035% for women versus 16.77% for men. The same goes for women filing patent applications together with another person, thereby forming a team of two patentees, compared to men filing a patent application with another person, with respectively 16.843% and 22.485%. Women filing patent applications in teams with more than five applicants are also more frequent than men filing patent applications in teams with more than five applicants, with respective shares of 32.039% and 28.759%. These observations indicate that, in general, women's inventorship share is lower than men's inventorship share.

4. Methodology

As previously mentioned, two different datasets were considered to investigate hypotheses H2, H3 and H4: a cross-sectional dataset observing all regions at one point in time as well as a panel data set observing all regions across time, from 2013 to 2017.

4.1 Cross-sectional data

4.1.1 Preliminary considerations

The methodological approach to be followed was a major point of concern in this research. Namely, the major problem was that the dependent variable considered is expressed as a percentage. Put differently, the dependent variable considered in this paper can be defined as the proportion of patent applicants in each NUTS2 European region that are women. Thus, the dependent variable can be considered as a proportion, and the appropriate model used should be chosen accordingly. The dependent variable is as such bounded between 0 and 1, meaning that the effect of explanatory variables tends to be non-linear, and that the variance tends to decrease when the mean gets closer to one of the boundaries. As a result, linear regression models are less appropriate when the dependent variable is a share, a percentage or a proportion (Buis, 2010). A more appropriate model is a beta regression, which is based on the assumption that the response is Beta distributed (Ferrari & Cribari-Neto, 2004). Smithson and Verkuilen (2006, p. 54) similarly argue that "for scales with a lower and upper bound, a suitable candidate for models is the beta distribution". Another strong argument in favour of Beta regressions is the fact that the support for the Beta distribution lies between 0 and 1, making Beta distributions a natural choice for modelling percentages (Swearingen, Melguizo Castro, & Bursac, 2011).

4.1.2 Beta regression

Preliminary explanations about the Beta regression model are required. This model is based on the assumption that the dependent variable follows a beta distribution, with a two-parameter distribution function as follows (Ferrari & Cribari-Neto, 2004):

$$\pi(y;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{(\alpha-1)}(1-y)^{\beta-1},$$

where $y \in (0,1)$, $\alpha > 0$, $\beta > 0$ and Γ is the gamma function. The mean and the variance are defined as follows, respectively:

$$E(y) = \frac{\alpha}{\alpha + \beta}$$

and

$$V(y) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

α and β are the two parameters indexing the distribution, i.e. its shape.

The model can also be defined including a dispersion parameter denoted ϕ and the mean of the response variable denoted μ. This new parametrisation writes the density function as follows:

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma\big((1-\mu)\phi\big)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1},\ y \in (0,1),$$

where $0 < \mu < 1$ and $\phi > 0$.

Let $\mu = \frac{\alpha}{\alpha+\beta}$ and $\phi = \alpha + \beta$. We now have:

$$E(y) = \mu$$

and

$$V(y) = \frac{\mu(1-\mu)}{1+\phi}$$

With the cross-sectional dataset used in this paper, the dependent variable y corresponds to the women share in patent applicants across 267 NUTS2 regions. The variable's distribution plot as well as the associated parameters α, β, μ and ϕ – as well as the variable's mean and variance – are presented in figure 2 and table 5 respectively[19]:

[19] Detailed calculations of the Beta-regression parameters can be found in appendix A.

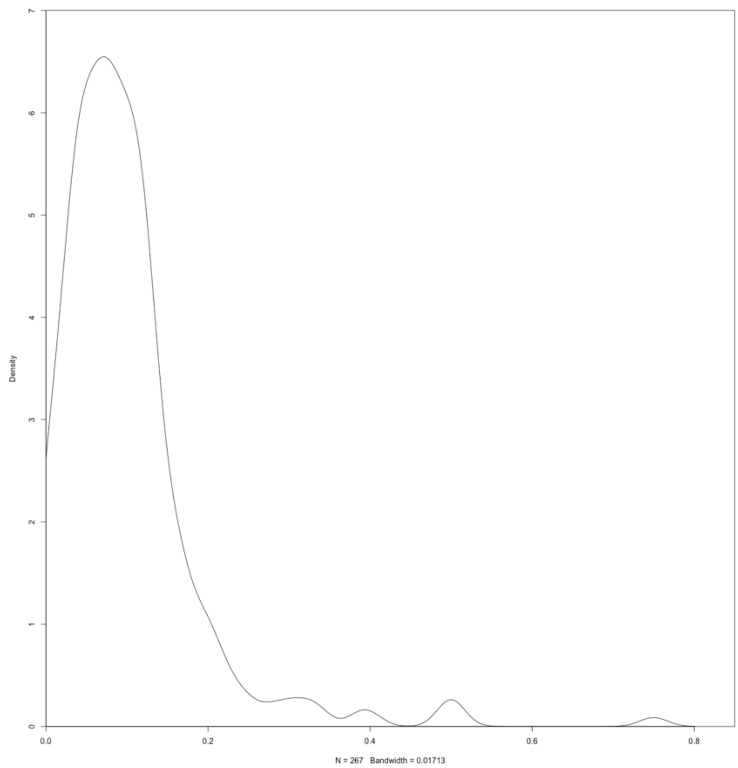


Figure 2. Dependent variable distribution

	Parameters	Value
1	α	1.017
2	β	9.207
3	μ	0.100
4	ϕ	10.224
5	E(y)	0.100
6	V(y)	0.008

Table 5. Beta-regression parameters

Further, the Beta-regression model was specified as follows:

$$g(\mu_i) = \beta_1 + \beta_2 x_{i2} + \cdots \beta_j x_{ij} + e$$

where $x_2, \dots, x_j$ stands for the set of explanatory variables (see table 2).

The link function used was a logit link, as follows:

$$\ln\left(\frac{\mu_i}{1-\mu_i}\right) = x_i\beta$$

The dependent variable contained a non-negligeable number of regions without any women patentee (18), resulting in the dependent variable taking the value zero for these regions[20]. In order for 0s to be considered in the estimation, the fem_share dependent variable was transformed using Smithson's and Verkuilen's proposition (2006, p. 61):

$$y_i' = \frac{[y_i(n-1) + 0.5]}{n}$$

Here, $n = 267$. For instance, let us consider region i for which fem_share variable is originally zero, i.e. $fem\_share_i = 0$. With the above-mentioned transformation, fem_share in region i takes value: $fem\_share_i' = \frac{0*(267-1)+0.5}{267} = \frac{1}{534} > 0$.

4.1.3 Robustness

A major threat to the robustness of estimated coefficients lies in the violation of the homoskedasticity assumption $V(\varepsilon_i|x_i) = \sigma^2$, i.e. the variance of the error term is the same across observations and does not depend on the value of the set of explanatory variables x_i. If this assumption does not hold, there is heteroskedasticity and standard errors robust to heteroskedasticity must be used. All regression models estimated with the above-detailed methods went through the Breusch-Pagan test[21], which has homoskedasticity for null hypothesis ($H0$). The null hypothesis was rejected anytime the p-value resulting from the Breusch-Pagan test was below 0.05, assuming the presence of heteroskedasticity. Heteroskedasticity was then addressed by using heteroskedasticity-robust standard errors. Furthermore, perfect multicollinearity was avoided by excluding variables similar to those presented in table 2 but considering the male gender. Obviously, including such variables would have resulted in perfect multicollinearity as the male variables are often the exact complementary to one of the female variables. Lastly, outliers were also controlled for by winsorizing the variables considered at the 5% level.

---

[20] 18 observations have no women patentee at all, i.e. 6.74% of observations included in the sample.

[21] See appendix B.

4.2 Panel Data

As mentioned before, this study also includes a panel dataset including observations for the same seventeen variables across the same 267 NUTS2 regions over the period from 2013 to 2017. Formally, the general panel data regression model is (McManus, 2011):

$$y_{it} = \beta_0 + \beta_k x_{kit} + \varepsilon_{it},$$

where: $i = 1, \dots, 267$ is the set of NUTS2 regions observed, $t = 2013, \dots, 2017$ is the time interval during which the observations are observed and $k = 1, \dots, 17$ is the set of variables observed for the 267 NUTS2 regions over the specified time interval.

4.2.1 Fixed effects

A panel regression model can also be formally written as follows:

$$y_{it} = \beta_0 + \beta_k x_{kit} + \beta_z z_i + \varepsilon_{it},$$

where z_i captures heterogeneities that do not change over time and are unobserved across the $i = 1, \dots, 267$ NUTS2 regions. Therefore, the omission of z_i could result in the presence of an omitted variable bias. The aim is again to estimate the set of β_k's, i.e. the effect of a change in the explanatory variables x_k on the dependent variable y_i, holding the unobserved and time-invariant factor z_i constant (Torres-Reyna, 2007).

Entity-fixed effects

The objective when applying an entity-fixed effect model is to account for and address an omitted variable bias that might arise due to variables that vary across regions but not over time. The entity-fixed effects regression model with a single regressor can be written as (Stock & Watson, 2020):

$$y_{it} = \alpha_i + \beta_1 x_{it} + \varepsilon_{it}$$

Generalising to k regressors, the model becomes:

$$y_{it} = \alpha_i + \beta_1 x_{1it} + \cdots + \beta_k x_{kit} + \varepsilon_{it},$$

where, $i = 1, \dots, 267$, $t = 2013, \dots, 2017$ and x_{kit} is the value of k^th^ regressor for entity i in year t. The specificity of this model is that it includes the term α_i. The terms $\alpha_i, \dots, \alpha_{267}$ represent entity-specific intercepts. We also have: $\alpha_i = \beta_0 + \beta_z z_i$.

Stock and Watson (2020) also explicit an equivalent way of writing the entity-fixed effect model, which includes a common intercept, the k regressors x_k and $n - 1$ binary variables that represent all but one region:

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \cdots + \beta_k x_{kit} + \gamma_2 D2_i + \gamma_3 D3_i + \cdots + \gamma_n Dn_i + \varepsilon_{it},$$

where $D2_i = 1$ if $i = 2$ and $D2_i = 0$ otherwise, and so on.

The first one of these two expressions is considered in the following sections, and the estimation method used is an entity-demeaned OLS[22] regression.

Time-fixed effects

Similarly, the objective when applying a time-fixed effect model is to account for and address an omitted variable bias that might arise due to variables that vary over time but not across regions. Stock and Watson (2020) indicate that the time-fixed effects regression model with a single regressor x can be written as:

$$y_{it} = \lambda_i + \beta_1 x_{it} + \varepsilon_{it}$$

Again, generalising the above expression to k regressors leads to:

$$y_{it} = \lambda_t + \beta_1 x_{1it} + \cdots + \beta_k x_{kit} + \varepsilon_{it},$$

where, again, $i = 1, \dots, 267$, $t = 2013, \dots, 2017$ and x_{kit} is the value of k^th^ regressor for entity i in year t. The specificity of this model is that it includes the term λ_t, which represents the effect on the dependent variable y of year t. The terms λ_t can therefore be defined as the time-fixed effects.

The time-fixed effect regression model can also be written with $T - 1$ binary variables:

[22] Ordinary Least Squares.

$$y_{it} = \beta_0 + \beta_1 x_{it} + \cdots + \beta_k x_{kit} + \delta_2 B2_i + \cdots + \delta_T BT_i + \varepsilon_{it},$$

where $B2_i = 1$ if $t = 2$ and $B2_i = 0$ otherwise, and so on.

The first one of these two expressions is considered in the following sections, and the estimation method used is a time-demeaned OLS regression.

Two-way fixed effects

The combination of entity – or region – and time-fixed effects models leads to a so-called two-way fixed-effect model, which addresses for omitted variable biases that might arise due to both variables that vary over time but not across regions and variables that vary across regions but not over time (Stock & Watson, 2020). The two-way fixed-effects regression model with k regressors can be written as:

$$y_{it} = \alpha_i + \lambda_t + \beta_1 x_{1it} + \cdots + \beta_k x_{kit} + \varepsilon_{it},$$

where $i = 1, \ldots, 267$, $t = 2013, \ldots, 2017$ and $k = 1, \ldots, 17$. Region and time-fixed effects are incorporated using the time and region demeaning method.

4.2.2 Robustness

A major point of concern regarding the robustness of the estimates calculated through the region, time and two-way fixed effects regression models is that it is impossible that observations within one region are independent over time. For example, considering that most education programmes last for more than one single year, whether a given NUTS2 region has a high share of female pupils attending an early childhood education programme in a given year is a good predictor of whether this region will have a high share of female pupils attending a similar programme the year after. This results in possible autocorrelation within regions observed, meaning that the error term ε_{it} for a specific region in a given year is likely to be correlated with its value in the subsequent year. To ensure the robustness of the models applied to the panel dataset, the standard errors estimated and used for inference were clustered standard errors. The clustered standard errors present a double advantage. First, they allow for autocorrelation within entities. Second, they are robust for heteroskedasticity (Cameron & Miller, 2015). Again, outliers were also controlled for by winsorizing the variables considered at the 5% level.

5. Empirical results

As previously mentioned, the results of this paper are split into two components: the identified geographical patterns used to test H1 and the results from the regression models explained in section 4 used to test H2, H3 and H4.

5.1 Geographical patterns

This section is about the first category of results, i.e. the geographical patterns. First, some preliminary observations are made. These first observations highlight the presence of a clustering of European countries into groups with two similar characteristics: their number of patentees and the share of women among their respective total number of patentees. Then, further and deeper observations are presented at the regional level, both within each cluster of countries and by country at the regional level. Using spatial visualisation on maps, a clearer pattern appears: the share of women among all patentees tends to be higher in regions where either the capital city or a large city of a given country is located, thereby confirming hypothesis H1 despite some non-negligible deviations. Moreover, several examples show that there may exist a spillovers effect affecting regions in the immediate vicinity of a neighbouring cluster.

5.1.1 Preliminary observations

Following the data collection procedure described in section 3, the number of patentees – regardless of their gender – in the selected countries from 2000 to 2018 is highest in Germany, second highest in France, and third highest in the United Kingdom[23]. Germany clearly has a strong preponderance in European patenting, as shown in figure 3. Moreover, the relatively low women share in Germany – third lowest among all countries considered – reduces the weighted average of women patentees share in all countries considered, which is 7.749% (see table 6).

[23] Respectively 53.47%, 11.16% and 6.36% of total; see figure 3.

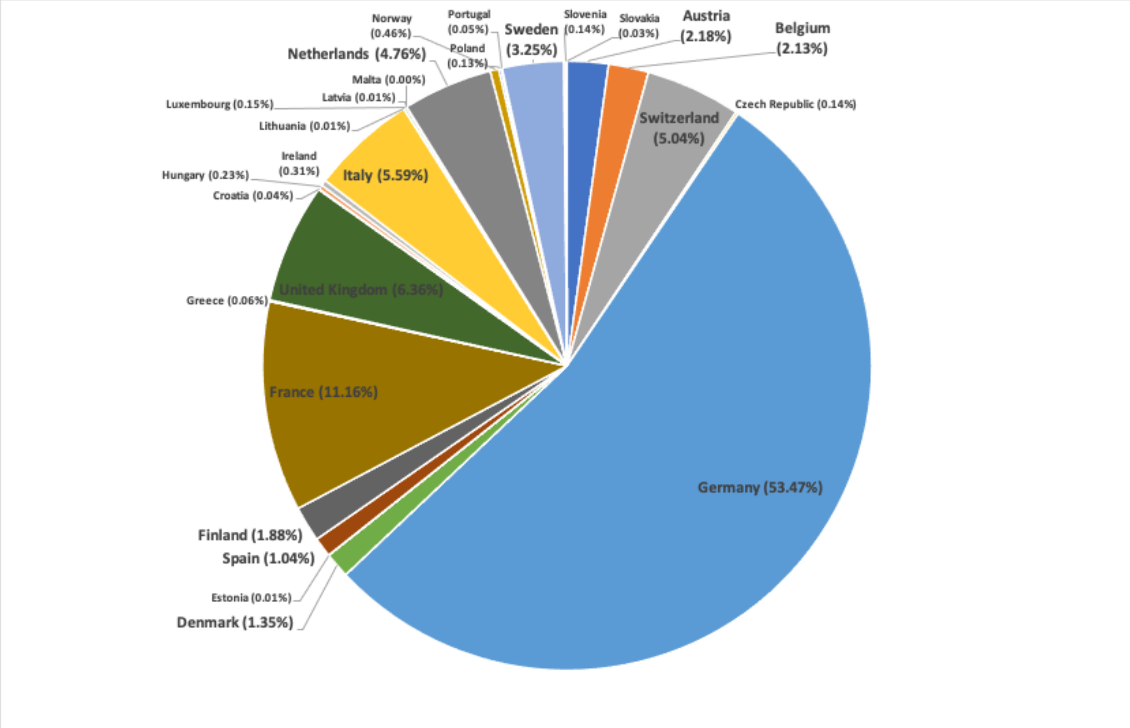


Figure 3. The weight of considered countries in the total number of patent applicants (2000-2018)

	Country	Patent Applicants	Women Share (%)
1	Austria	15, 376	4.188
2	Belgium	14, 989	11.303
3	Croatia	257	21.401
4	Czech Republic	977	6.448
5	Denmark	9, 352	9.859
6	Estonia	57	10.526
7	Finland	13, 241	13.843
8	France	78, 648	13.240
9	Germany	376, 668	5.590
10	Greece	454	12.115
11	Hungary	1, 608	21.580
12	Ireland	2, 176	9.191
13	Italy	39, 378	13.175
14	Latvia	90	33.333
15	Lithuania	77	11.688
16	Luxembourg	1, 076	6.134
17	Malta	25	4
18	Netherlands	33, 545	6.579
19	Norway	3, 213	11.360
20	Poland	947	15.628
21	Portugal	353	12.181
22	Slovakia	191	8.901
23	Slovenia	998	27.555
24	Spain	7, 319	15.945
25	Sweden	22, 926	8.314
26	Switzerland	35, 513	6.130
27	United Kingdom	44, 812	8.112
28	Total	704, 266	7.749

Table 6. Number of patent applicants and women share per country, 2000-2018

Based on the previous observations on countries' women share among their patentees, it is possible to cluster countries into groups, as shown in figure 4. The countries clustered into the same group present two similarities: a similar share of women among their patentees as well as a similar number of patent applicants, taken in its logarithmic transformation. The logarithmic transformation enables a clear visualisation of this phenomenon, addressing for the preponderance of Germany that would otherwise make a clear visualisation impossible.

Figure 4. Country clustering

Table 7 hereunder summarises the characteristics of each country group or cluster[24]. The central European group clearly stands out, both representing the highest share of the total sample and having the lowest share of women patentees.

[24] Colour code for clusters: Central Europe = red, Western Europe = yellow, Northern Europe = dark blue, Southern and Eastern Europe = green, Eastern Europe = light blue

	Cluster	Average Women Share, Weighted (%)	Weight in Total Sample (%)
1	Central Europe	5.660	65.760
2	Western Europe	13.160	19.930
3	Northern Europe	9.260	13.590
4	Southern and Eastern Europe	13.370	0.300
5	Eastern Europe	23.940	0.420

Table 7. Clusters' characteristics

The spatial visualisation of data– namely, the women share in patentees by NUTS2 region – in the subsequent section provides further insights into distinguishable geographical patterns.

5.1.2 Cluster and country level observations

The visualisation of each cluster on maps – figures 5, 6 and 9 hereunder – gives further details about the discrepancies in the women share across countries within clusters and provides first indications about the discrepancies in the women share across regions within the same country. Considering their small weight in the total sample, maps for the southern and eastern European as well as the eastern European clusters are not included, although available in appendix C.

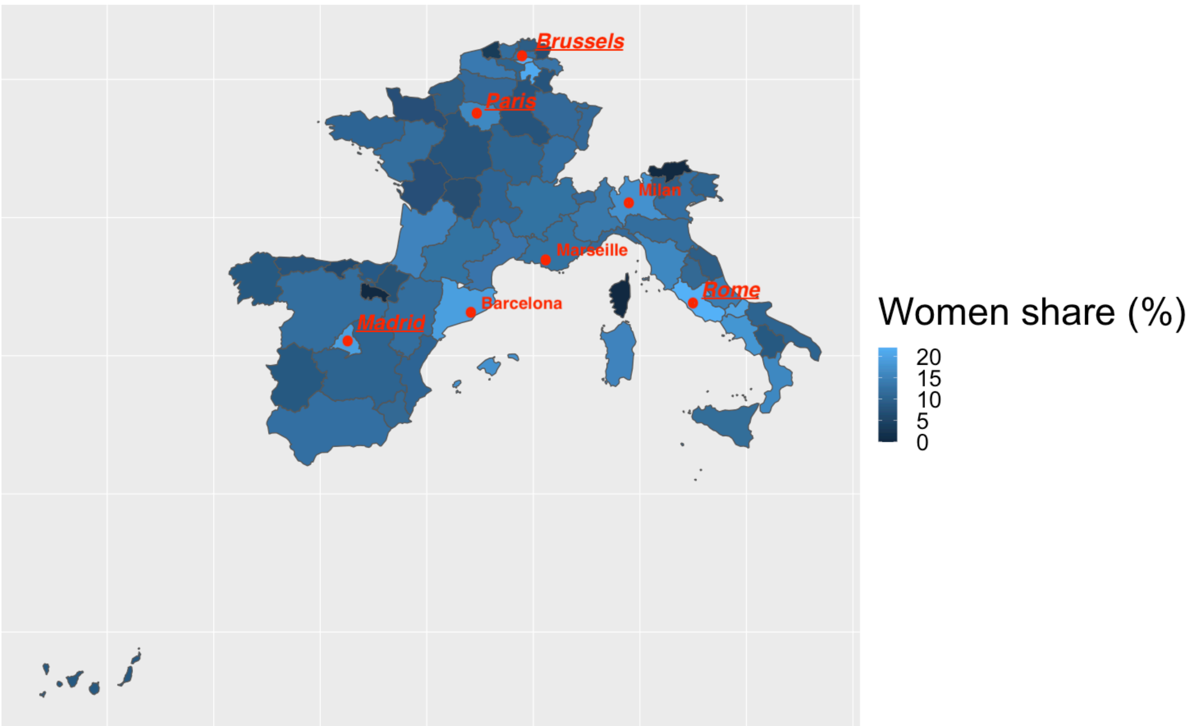


Figure 5. Women share within the western European cluster

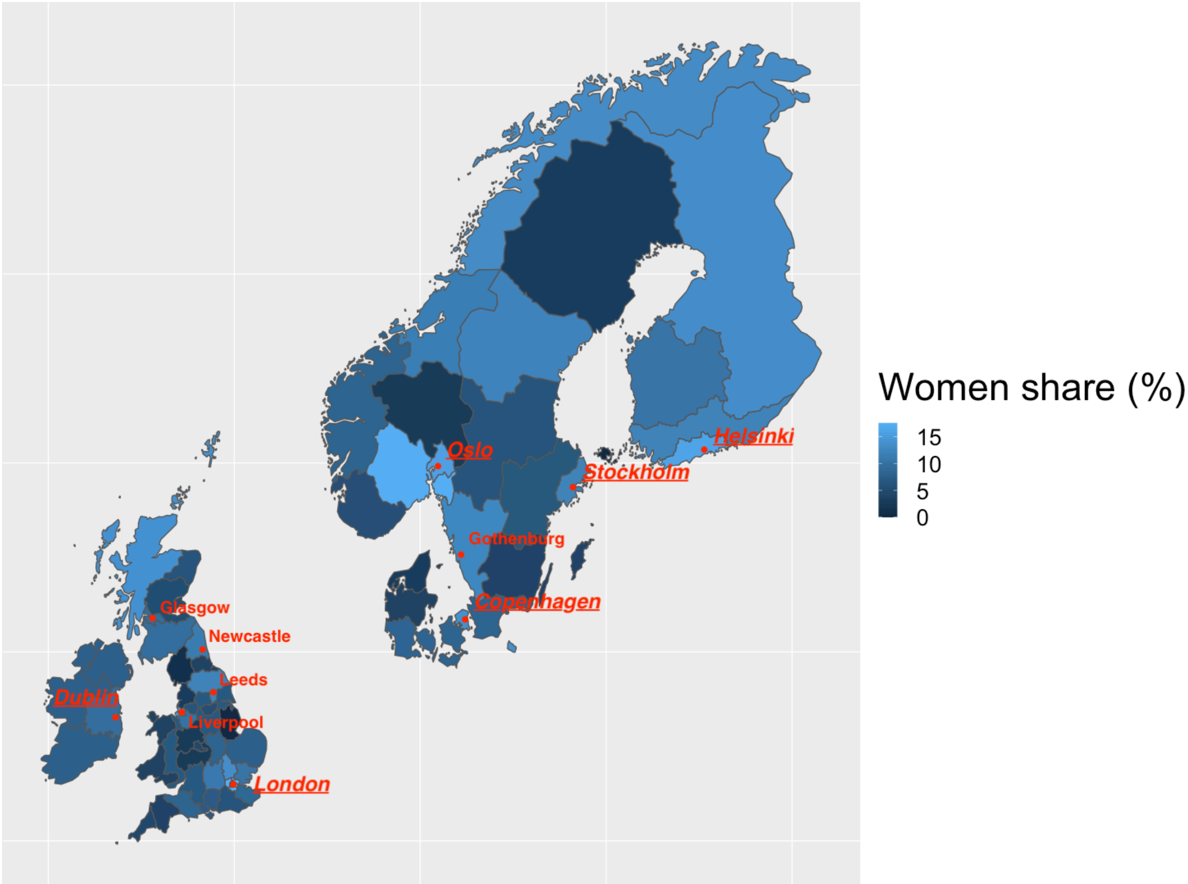


Figure 6. Women share within the northern European cluster

As mentioned in the hypotheses section, one of the initial conjectures was that the share of women among patentees is relatively higher in the NUTS2 regions where a large city is located, compared to the NUTS2 regions without a large city. This pattern is broadly confirmed by the observation of figures 5 and 6. Indeed, for these clusters, the women share of patentees tends to be higher in regions that contain either a capital or large city. Some examples of this pattern can be seen in France where Paris stands out and has the highest share of women patentees among the French regions, in Spain where Madrid and Barcelona have the highest women shares in the country, in Italy with Rome and Milan, and the United Kingdom with London, Leeds, Liverpool and Newcastle. The same pattern appears on figure 6 in Norway, Sweden, Finland, Denmark and to a lesser extent in Ireland[25].

[25] See appendix D for the individual maps of all countries.

However, two interesting cases are worth of notice: Belgium and the Italian northernmost region of South Tyrol[26]. Firstly, the highest women share in Belgium is not located in the country's capital city – Brussels – but in the surroundings of Charleroi (see figure 7), which can nonetheless be considered as a large city at the Belgian scale. Moreover, the country seems to be split, just as it is from the political and cultural perspectives. The northern Flemish-speaking part of the country – Flanders – shows a relatively lower women share compared to the southern French-speaking part of the country[27] – Wallonia – as shown on figure 8. Secondly, South Tyrol is the Italian region with the lowest share of women patentees, as shown on figure 8. This region presents the particularity of having a majority of German native speakers among its population, accounting for approximately two thirds of the regional population[28]. As such, the region is granted a large autonomy relating to its political, economic, educational and linguistic systems since 1948. Not only linguistically but also historically, the region has close ties to neighbouring Austria, to such an extent that both the provisional Austrian government and the South Tyrolese representatives intensely claimed and advocated for a return of South Tyrol to Austria immediately after the end of World War II[29] (Peterlini, 2007). Therefore, and in many aspects, South Tyrol can be considered very close to Austria, despite being an integral part of Italy. This multifaceted proximity was materialised in the creation of the Tyrol-South Tyrol-Trentino Euroregion in 1998, one aim of Euroregions being the support of business cooperation across borders (Durà et. al, 2018), thereby reinforcing the ties between South Tyrol and neighbouring Austria.

[26] Also known as Alto Adige in Italian and Südtirol in German.

[27] Flemish regions have an average women share of 8.65% while the Walloon regions have an average women share of 14.77%.

[28] According to the Provincial Institute of Statistics (ASTAT: *Istituto Provinciale di Statistica*), based on the 2011 census (see: https://astat.provincia.bz.it/downloads/JB2014.pdf)

[29] South Tyrol was annexed by Italy in 1919 following the Treaty of Saint-Germain, in the aftermath of World War I. Before 1919, South Tyrol was part of Austria-Hungary double monarchy under the Austrian crown.

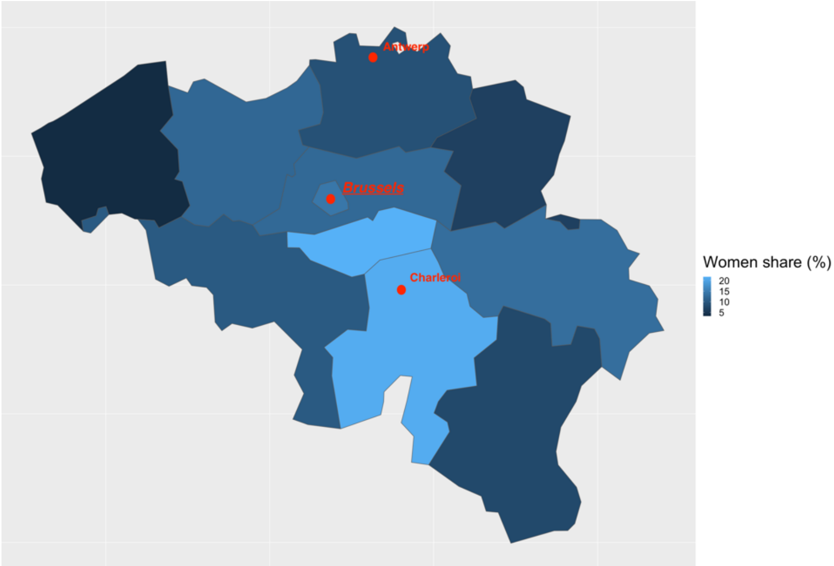


Figure 7. Women share in Belgium

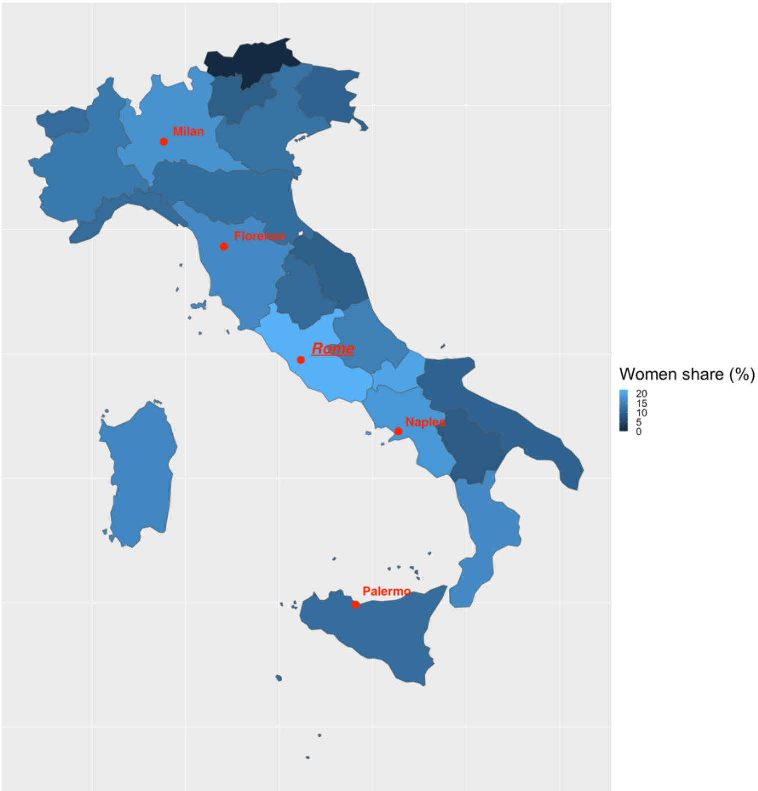


Figure 8. Women share in Italy

Noticeably, Flanders and South Tyrol can be seen as peripheric regions of the western European cluster, due not only to their geographic proximity but also to their close cultural ties to countries other than the country they are part of: Flanders maintains close ties with the neighbouring Netherlands while South Tyrol has close ties with Austria, both the Netherlands and Austria belonging to the central European cluster. This leads to a closer observation of this cluster to see whether a similar pattern exists there.

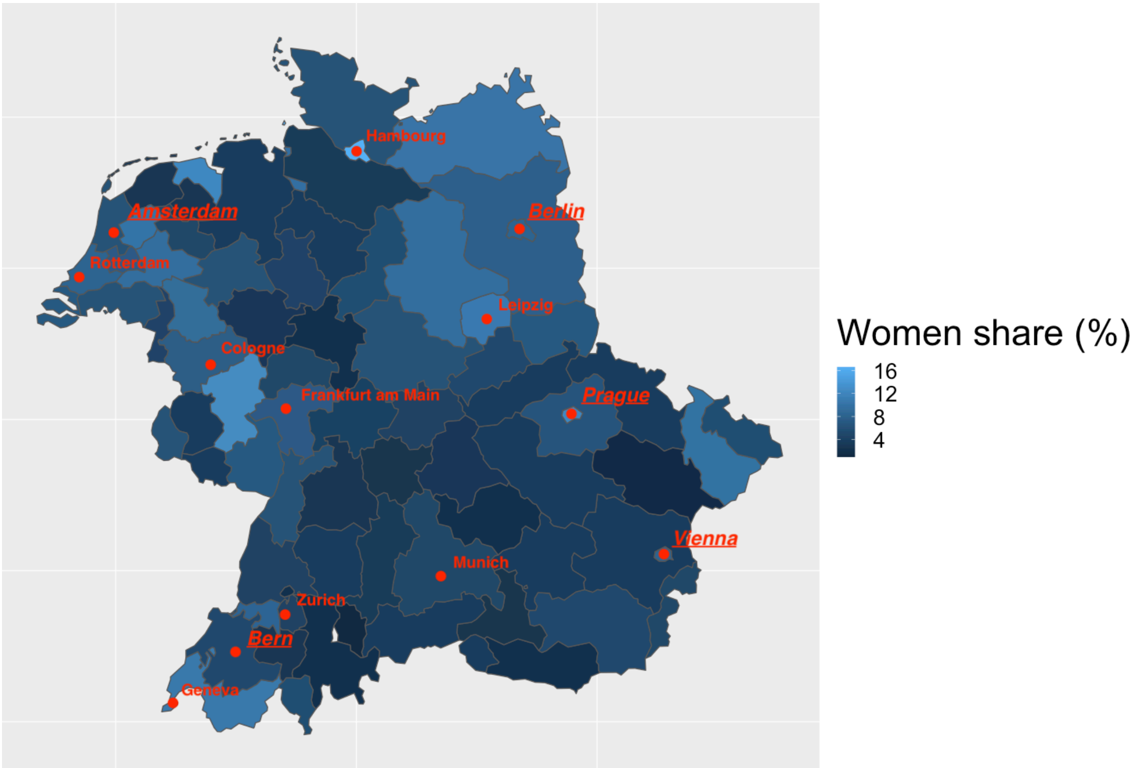


Figure 9. Women share within the central European cluster

Dominated by Germany, the central European cluster does not give such clear indication whether the city-effect hypothesis is valid or not. Indeed, some large cities like Hamburg, Prague, Vienna, Geneva or Leipzig present a relatively high share of women patentees in their respective countries, but most other large cities do not seem to have a stimulating effect on women patenting in their respective region (see figure 9).

In a similar way as for South Tyrol and Flanders, another particular case is to be observed in Switzerland within the central European cluster. The French-speaking peripheric region of Geneva bordering France presents an unusually high share of women patentees compared to other Swiss regions. Again, a Euroregion exists in this area since 1974, which can be beneficial in a context of globalised and intensified competition (Sohn, Reitel, & Walther, 2009).

5.1.3 City effect confirmation, central European specificity and cross-border spillovers

Based on the observations made in the two previous sections, three sets of insights can be drawn. They first tend to confirm the city effect, although some non-negligible deviations exist. Furthermore, the existence of special cases in regions located at the border between two countries belonging to different clusters – central European and western European – advocates for the existence of cross-border influence between countries, that is cultural spillovers. Finally, the central European cluster, and perhaps Germany even more, present employment and historical features that certainly play a role in the women patenting activity.

Mild confirmation of the city effect

The first hypothesis, stated in section 2.1, was that the share of women among the total patentee population increases around a given country's large cities. A majority of the European countries present their highest shares of women patentees in a region that contains a large city, on each country's respective scale[30]. There are four exceptions: Croatia[31], Hungary[32], the Netherlands[33] and Poland[34].

However, the city effect is much less observable in the central European cluster, notably in Germany (figure 10), the Netherlands (figure 11) and Switzerland (figure 12). Germany includes two industrial and economic hubs with a high employment density along the river Rhine, in the western and north-western parts of the country: the Ruhr area and to a lesser degree the Frankfurt-Mainz area. Despite the fact that a high employment density is a strong driver of patenting activity (Carlino, Chatterjee, & Hunt, 2007), the high density in these two German regions does not translate into a higher share of women patentees. Similarly, the large cities of Munich and Cologne do not seem to stimulate women patenting in their respective region. The Swiss region with the highest women share is the Geneva region, undeniably a large city in Swiss terms. However, the two other large Swiss cities – Zurich and Bern, the capital city – present a relatively low women share among patentees. Nonetheless, Vienna and

[30] The following countries have only one NUTS2 region, and it is therefore impossible to validate or not this hypothesis in their case: Estonia, Latvia and Luxembourg.

[31] The Zagreb region has a women share of 19.2% compared to 23.5% in the country's other region located in the Adriatic region. See appendix D.

[32] The Budapest region has a women share of 21.1% compared to 23.2% – highest value in the country – in the region of the Northern Great Plain. See appendix D.

[33] The regions of Amsterdam and Rotterdam present a women share of 6% and 8.2%, respectively, while the country's highest value is of 12.5% in the Groningen province. See appendix D and figure 11.

[34] The Warmia-Masuria province has a women share of 20.3% – highest in the country – while the Polish provinces containing large cities all have a women share lower than 20%. See appendix D.

Prague are the regions with highest women share in Austria and Czech Republic, respectively (figures 14 and 15).

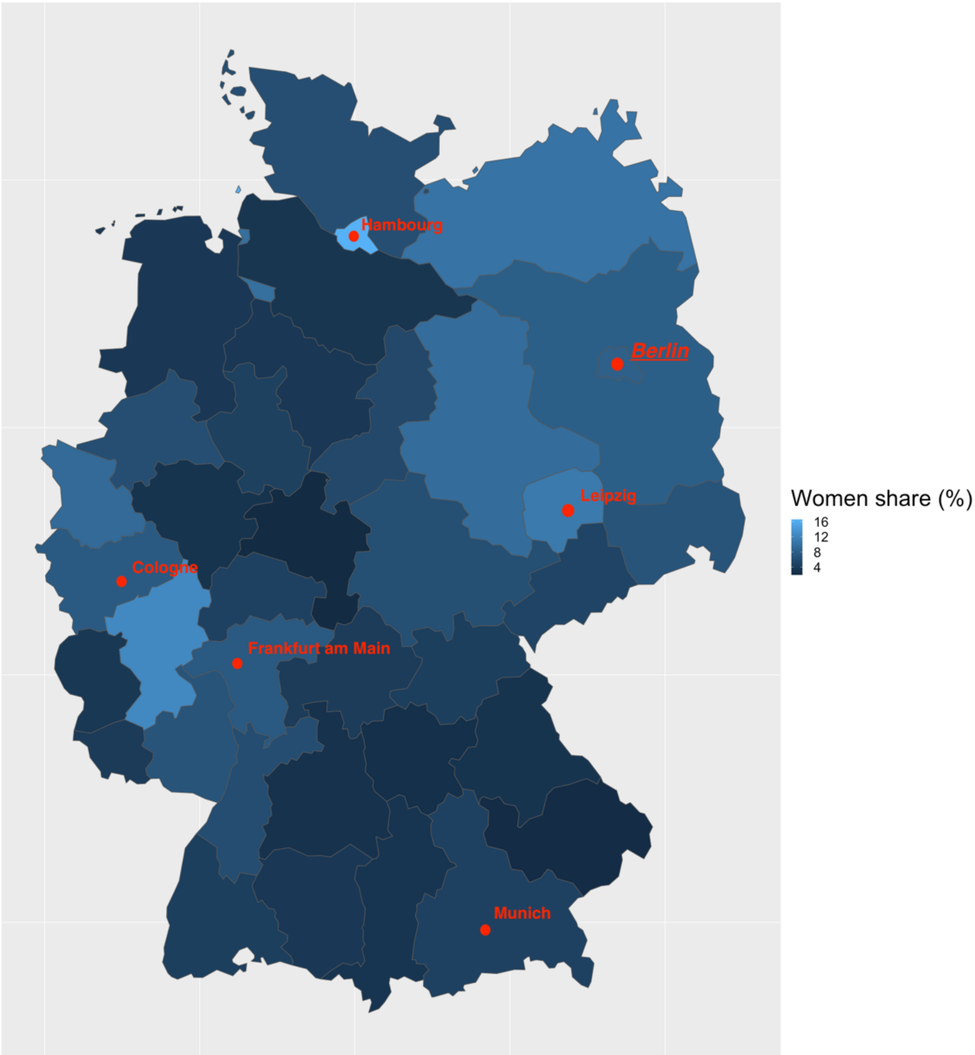


Figure 10. Women share in Germany.

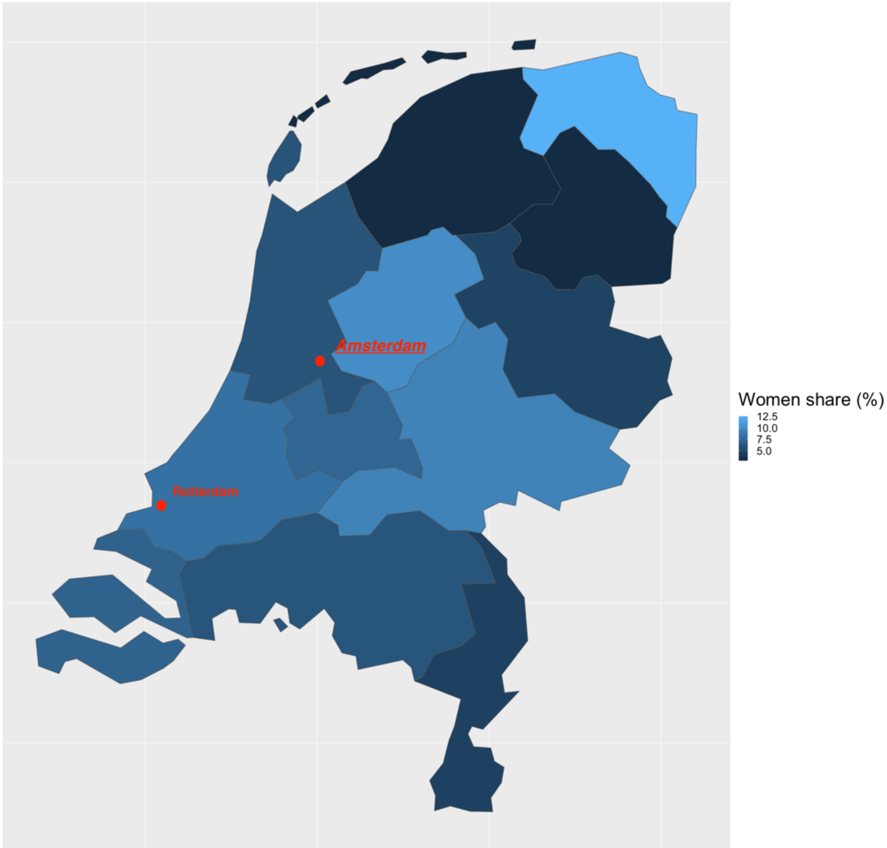


Figure 11. Women share in the Netherlands.

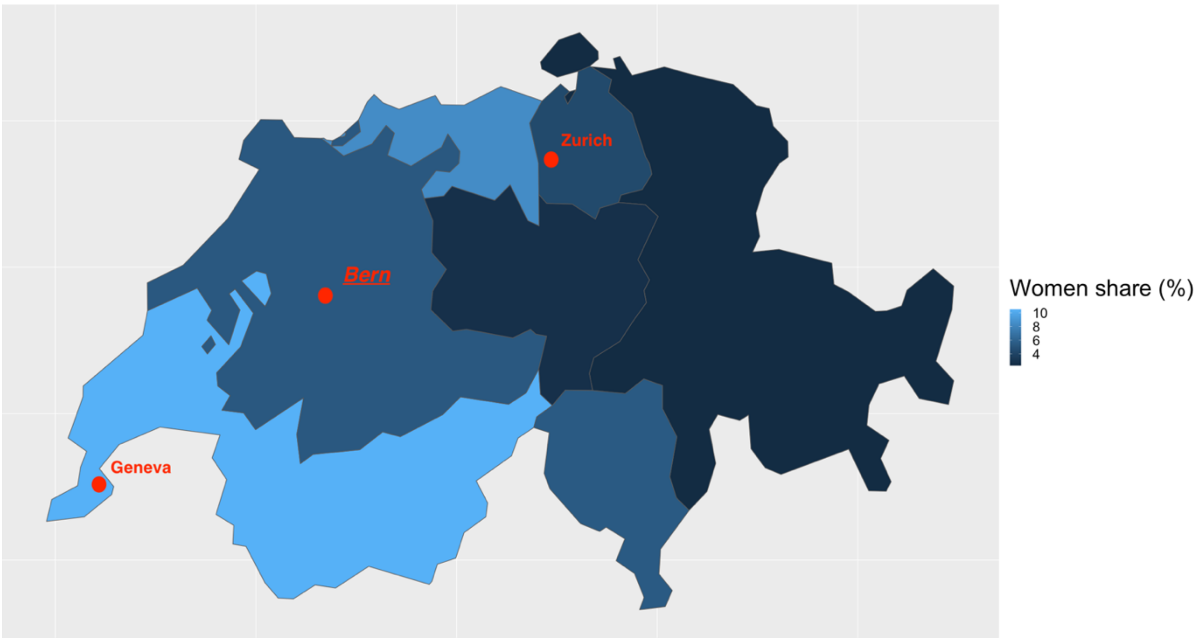


Figure 12. Women share in Switzerland

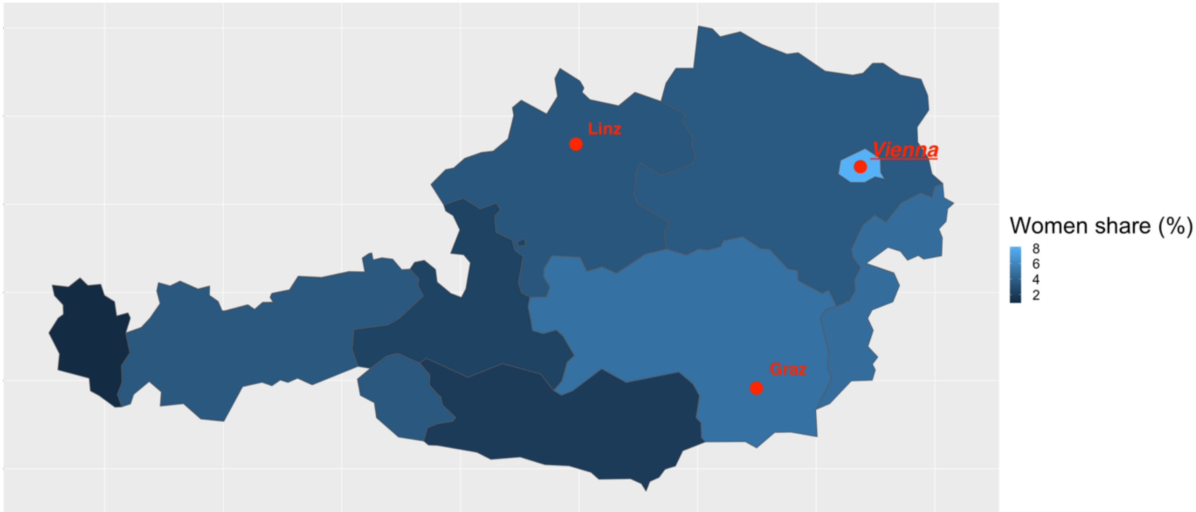


Figure 13. Women share in Austria

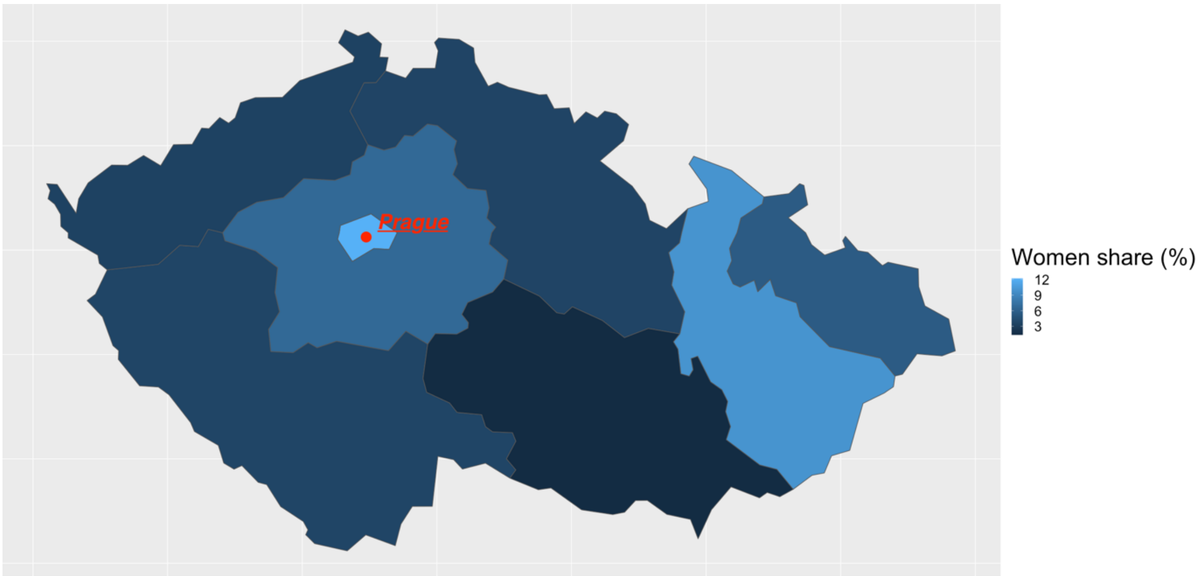


Figure 14. Women share in Czech Republic

There are strong signs of the existence of a city effect stimulating patenting by women in the western and northern European clusters, while the central European cluster – in particular Germany, Switzerland and the Netherlands – casts a doubt on its existence. Whether the city effect does not exist or the central European cluster is an exception remains to be seen – this question is discussed in the immediately following section.

The central European specificity

As previously mentioned, there are several instances in the central European cluster that clash with the city-effect hypothesis. These instances, however, do not necessarily call the city effect into question, as there are two possible explanations for these contrasting regions and countries.

Gender gap in part-time employment

The first plausible explanation relates to the gender gap in part-time employment in the countries in question, namely Germany, the Netherlands, Switzerland and Austria. This gender gap is defined as the difference between the share of part-time employment in total employment of women and men aged 20 to 64[35]. In these countries, the gender gap in-part time employment is the most severe in Europe, meaning that much more German, Dutch, Swiss and Austrian women work part time compared to their counterparts in other European countries, putting them at a disadvantage in many terms, including productivity (Matteazzi, Pailhé, & Solaz, 2017). This aligns with Stephan's and El-Ganainy's (2006) statement about women having a lower productivity, which in turn can put women at a disadvantage in many situations and slow down their patenting activity. This could explain not only why the city effect seems less effective in the central European cluster, but also why the average women share in these countries is much lower than the sample average excluding these four countries[36].

Year	Netherlands	Switzerland	Germany	Austria	EU28
2017	51.5	46.0	37.5	37.4	23.0
2016	52.7	46.7	37.9	37.4	23.2
2015	52.9	47.0	38.0	38.1	23.3
2014	53.0	47.8	37.8	37.8	23.5
2013	53.5	48.8	38.3	37.0	23.7

Figures expressed in %

Table 8. The part-time employment gender gap in central Europe and the EU

Moreover, a likely reason for German women taking more part-time employment is motherhood, and German working mothers being in part-time employment in a much higher proportion than German working fathers (Weinkopf, 2014), deteriorating women's work-life

[35] See Eurostat: https://ec.europa.eu/eurostat/databrowser/view/tepsr_lm210/default/table?lang=en

[36] The sample weighted average, excluding the Netherlands, Germany, Austria and Switzerland is a women share of 11.70%, while the weighted average women share in these four countries is only 5.66%.

balance conditions, especially with dependent children to take care of (Wynarczyk & Renner, 2006). This feature likely holds in Austria, the Netherlands and Austria as well.

The Soviet legacy

The second plausible explanation lies in history, which can explain two observations. Firstly, the former separation between eastern and western Germany seems to be somehow still present when considering the share of women in the total patentee population. Indeed, by a simple map observation, the German regions located in the former Eastern Germany tend to have a slightly higher women share than the German regions located in the former Western Germany. Secondly, most of the other countries belonging to the former Soviet bloc – or communist Yugoslavia – have a relatively high women share compared to the other countries in the sample. This is the case for Croatia, Estonia, Hungary, Latvia, Lithuania, Poland and Slovenia[37]. These two observations are perhaps not surprising considering the important legacy of the former socialist and communist regimes in terms gender parity in the scientific fields (UNESCO, 2016).

Considering both the important part-time employment gender gap in the central European cluster as well as the soviet legacy in the eastern German regions, there are signs indicating the existence of a city effect, in the sense that women tend to represent a more important share of the patenting population in the large cities, or at least in their surroundings. Furthermore, the larger part-time employment gender gap in the central European cluster and its negative consequences on women's productivity and patenting participation gives some credit to the fourth hypothesis – that gender inequality negatively impacts women's patenting activity – although it remains to be confirmed.

Cross-border cultural spillovers

Borders between European states are no longer hard borders, materialised by the freedom of movement across most European countries and the introduction of several Euroregions. As previously mentioned, Euroregions aim at reinforcing the economic integration of two neighbouring regions that have close economic, historical and cultural ties but are however located on different sizes of an international border. Three examples of such Euroregions can

[37] Czech Republic and Slovakia are exceptions and were quite interestingly formerly united as Czechoslovakia both before and during the communist era.

be found between Italy and Austria in the South Tyrol region, between France and Switzerland in the Geneva area and between Belgium and the Netherland with the Belgian Flanders region.

Based on map observations, these three regions do not present the same characteristics as the other regions of the country and cluster they belong to, in terms of patenting activity by women. Indeed, and without clustering, and without marking national borders, Geneva has a share of women patentees that is similar to a French region. The same applies to the Italian South Tyrol and Austria, and to Belgian Flanders and the Netherlands. Hence, there are signs that cultural spillovers across international borders exist, and that these spillovers influence the patenting activity by women. Furthermore, the fact that each of these three regions considered is located at the border between two clusters indicates that these spillovers can relate to culture, as these regions present characteristics of the neighbouring cluster rather than the cluster encompassing the country they belong to. Therefore, it is likely that culture influences the frequency of women patenting, as all three South Tyrol, Flanders and Geneva are somehow culturally distinct from the rest of their respective country.

After investigating the validity of the first hypothesis and having identified some specific features of the central European countries based on the analysis of maps, the following subsection focuses on econometric empirical results. These results will be used to see whether there is support validating the three other hypotheses: a higher educational attainment level positively correlating with a higher patenting by women (H2), STEM academic disciplines further increasing women’s patenting (H3) and severe gender inequality putting women at a disadvantage when it comes to patenting (H4).

5.2 Gender gaps: educational attainment, STEM and gender inequality

5.2.1 H2 testing: cross-sectional and panel analyses

Cross-sectional analysis

	Dependent variable:			
	betatrans_fem_share			
	(1)	(2)	(3)	(4)
F_early_educ_count	0.00001***		0.00001***	
	(0.0000)		(0.0000)	
log(F_early_educ_count)		0.395***		0.318***
		(0.070)		(0.070)
F_share_tereduc_Bach	−0.035**	−0.033**	−0.032**	−0.030**
	(0.014)	(0.014)	(0.015)	(0.015)
F_share_tereduc_Mast	−0.009	−0.005	0.008	0.008
	(0.013)	(0.013)	(0.014)	(0.014)
F_share_tereduc_PhD	0.050***	0.045***	0.044***	0.041***
	(0.012)	(0.012)	(0.012)	(0.012)
F_share_pop			0.071	0.073
			(0.065)	(0.064)
F_EDat02			−2.099*	−2.010*
			(1.163)	(1.141)
F_EDat34			−1.977*	−-1.910*
			(1.159)	(1.137)
F_EDat38			−0.131***	−0.111***
			(0.041)	(0.040)
F_EDat58			−1.963*	−1.896*
			(1.159)	(1.137)
Constant	−2.425***	−6.286***	203.299*	191.495*
	(0.825)	(1.062)	(116.024)	(113.748)
Observations	139	139	139	139
R^2	0.163	0.237	0.272	0.317
Log Likehood	199.046	206.453	209.361	214.050
Note:		*p<0.1; **p<0.05; ***p<0.01		

Table 9. Beta regressions on educational gender gaps variables

Table 9 shows the regression results on educational attainment gender gaps variable, applying a beta regression method. The variables labelled F_EDat[38] and F_share_pop[39] are meant as control variables. The models include three variables capturing the gender gap in the participation in educational programmes at four different levels: early educational

[38] Share of women with education up to secondary, post-secondary and tertiary education.

[39] Share of regional population that are women.

programmes[40] as well as all three levels of higher education from bachelor to PhD[41]. A logarithmic transformation is applied to the variable F_early_educ_count in order to account for its likely non-linear relation with the dependent variable. Values reported between parentheses are standard errors of the coefficient estimates.

Specifications (2) and (4) give first indications about hypothesis H2 – the educational attainment level. The positive and statistically significant – at the 1% level – estimated coefficients on the variable F_share_tereduc_PhD indicate that the higher the share of women undertaking PhD programmes, the higher women's patenting activity is. The estimated coefficient on the F_share_tereduc_PhD is equal to 0.045 in specification (2) and to 0.041 in specification (4)[42]. The respective R^2 for specifications (2) and (4) are 0.237 and 0.317, i.e. respectively 23.7% and 31.7% of the variance in the share of women among patentees is explained by the independent variables. Since a logit link function was used in these specifications, these coefficients are log-odds. Hence, the coefficients estimated by such a beta regression correspond to the additional increase or decrease in the log-odds of the dependent variable. Formally, the model equation is the same as in a logistic regression, which corresponds to:

$$logit(\mu_i) = x_i\beta_i$$

$$\ln\left(\frac{E(y_i)}{1 - E(y_i)}\right) = x_i\beta_i$$

Where $\mu_i = E(y_i)$. In this setup, $y_i = fem\_share_i$, i.e. the dependent variable being defined as the share or proportion of women among the total patenting population. Formally, the dependent variable fem_share is defined as:

$$fem\_share_i = \frac{women\ patenting_i}{women\ patenting_i + men\ patenting_i}, \text{ in region } i.$$

[40] F_early_educ_count and its logarithmic transformation.

[41] F_share_tereduc_Bach at the bachelor level, F_share_tereduc_Mast at the master level, F_share_tereduc_PhD at the PhD level.

[42] Both estimates are statistically significant at the 1% level (p-value<0.01).

From specification (2), the estimated coefficient for the independent variable F_share_tereduc_PhD is significant at the 1% level and equal to 0.045. Hence, an absolute 1-unit change in F_share_tereduc_PhD leads to a relative change of $e^{0.045} \approx 1.046$, that is a 4.6% positive change in the proportion of women in patenting, thereby reducing the patenting gender gap in region i. Considering specification (4), the change in the proportion of women in patenting is also positive, and equal to $e^{0.041} = 1.0419$, that is a 4.19% increase[43]. Hence, the higher the share of women in the PhD population, i.e. the narrower the PhD gender gap, the higher the proportion of women among the patenting population, i.e. the narrower the patent gender gap. Interestingly, the estimated effect of narrowing the gender gap at the bachelor level has an opposite sign. Indeed, from specification (2) and (4), the estimated coefficient for the F_share_tereduc_Bach variable is equal to -0.033 and -0.030, respectively, both estimates being statistically significant at the 5% level. Considering specification (4), the change in the proportion of women in patenting is negative and equal to $e^{-0.030} \approx 0.9704$, i.e. a 2.96% reduction. The estimated effect of the participation of young girls to early education programmes[44] on the patent gender gap is positive, as estimates from both specifications (2) and (4) are positive and significant at the 1% level. However, the estimated effect of narrowing the gender at the master level – captured by the variable F_share_tereduc_Mast – is not significant.

Overall, there is some evidence that a narrower gender gap at the PhD level – correlates positively with the proportion of women in the patenting population. However, the estimated negative correlation between the bachelor gender gap at the bachelor level and the patent gender gap casts doubts on the progressive nature of the improvements in the patenting gender gap as the gender gap in education levels narrows.

Panel analysis

The advantage of panel analyses is to make possible to address for omitted variable biases that may arise due to variables that vary over time and across regions when controlling for both time and entity-fixed effects, i.e. two-way fixed effects. As before, the F_EDat variables are considered as control variables.

[43] The coefficient estimate is significant at the 1% level.

[44] F_early_educ_count variable taken in its logarithmic transformation.

	Dependent variable:			
	fem_share			
	(1)	(2)	(3)	(4)
F_early_educ_count	0.001		0.001	
	(0.001)		(0.0005)	
log(F_early_educ_count)		23.413		26.576
		(21.425)		(21.322)
F_share_tereduc_Bach	0.464	0.574	0.478	0.562
	(1.069)	(1.097)	(1.100)	(1.118)
F_share_tereduc_Mast	−0.139	−0.060	−0.055	0.072
	(0.662)	(0.626)	(0.756)	(0.698)
F_share_tereduc_PhD	−0.622***	−0.622***	−0.665***	−0.642***
	(0.223)	(0.212)	(0.227)	(0.236)
F_share_pop			4.452	4.614
			(7.119)	(7.683)
F_EDat02			−7.538**	−8.537**
			(2.994)	(3.379)
F_EDat34			−7.443**	−8.345**
			(2.899)	(3.318)
F_EDat38			−6.878**	−7.012**
			(2.654)	(2.877)
F_EDat58			−7.567**	−8.545**
			(2.954)	(3.331)
Observations	188	188	188	188
R^2	0.073	0.063	0.110	0.113
Adjusted R^2	−0.354	−0.369	−0.342	−0.338
F Statistic	2.531** (df = 4; 128)	2.161** (df = 4; 128)	1.921** (df = 8; 124)	1.970** (df = 8; 124)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 10. Panel regressions, two-way fixed effects

Controlling for both time and entity fixed effects, the estimated coefficient on the PhD-level gender gap variable[45] is still highly significant – at the 1% level – but its sign has changed from positive (see table 9) to negative (see table 10). Put differently, narrowing the PhD gender gap does not seem to narrow the patenting gender gap, but instead to increase it. At the same time, the coefficient estimates on gender gaps at the other educational levels have lost their significance. The change in coefficient sign can be due to the fact that, so far, the regression models considered have not made any sort of distinction between the academic disciplines considered, and considered all kinds of PhD degrees, regardless of the field of study. Hence, it can be the case that, even if the gender gap narrows at the PhD level, there is no improvement in the patent gender gap due to the fact that women can very well attend more PhDs in non-scientific fields for which the likelihood of patenting is small to very small, like humanities.

[45] Denoted F_share_tereduc_PhD

This leads to the investigation of hypothesis (3), namely whether the academic discipline of degrees obtained impacts the severity of the patent gender gap.

5.2.2 H3 & H4 testings: STEM gender gap and gender inequality

	Dependent variable:						
	betatransfem_share						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
F_early_educ_count	0.0001***		0.0001***		0.0001***		
	(0.00000)		(0.00000)		(0.00000)		
log(F_early_educ_count)		0.436***		0.450***		0.423***	0.416***
		(0.082)		(0.085)		(0.082)	(0.087)
F_share_tereduc_Bach	−0.025	−0.025	−0.035**	−0.033**	−0.033**	−0.032**	−0.030**
	(0.016)	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)	(0.013)
F_share_tereduc_Mast	−0.011	−0.008	−0.003	0.001	−0.014	−0.010	−0.019
	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)	(0.023)
F_share_tereduc_PhD	0.044***	0.043***	0.037***	0.034***	0.047***	0.045***	0.044***
	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)	(0.012)	(0.011)
F_share_pop	0.067	0.072	0.103	0.118	0.027	0.063	0.072
	(0.083)	(0.082)	(0.095)	(0.094)	(0.098)	(0.096)	(0.101)
Share_Grad_STEM	−0.033**	−0.026*					−0.039**
	(0.017)	(0.017)					(0.018)
F_Share_STEM	0.023**	0.021**					0.019**
	(0.0011)	(0.011)					(0.010)
KnowlIntEmp			−0.009	−0.009			−0.014
			(0.012)	(0.012)			(0.016)
FemEmpAdvDeg			0.017	0.014			0.023
			(0.016)	(0.016)			(0.025)
GII					−3.364**	−3.642**	−3.599**
					(1.626)	(1.593)	(1.577)
GInnIx					−0.017	−0.015	−0.017
					(0.012)	(0.012)	(0.014)
Constant	−5.465	−9.887**	−8.261	−13.249**	−2.528	−8.423	−8.188**
	(4.303)	(4.204)	(5.208)	(5.157)	(5.394)	(5.305)	(5.237)
Observations	139	139	139	139	139	139	139
R^2	0.218	0.251	0.213	0.249	0.215	0.253	0.301
Log Likelihood	205.137	207.557	204.593	207.509	205.198	208.672	210.552
Note:					*p<0.1; **p<0.05; ***p<0.01		

Table 11. Beta regressions with STEM and inequality variables

As seen on table 11, the estimated coefficient on the PhD-level gender gap[46] is still highly statistically significant in all specifications, and positive. Considering specification (7), the estimated coefficient is equal to 0.044, which means that the estimated effect on the patent gender gap[47] is equal to $e^{0.044} = 1.0450$, i.e. a 4.5% positive change. This is very similar to results derived from the analysis of table 9.

[46] F_share_tereduc_PhD variable.

[47] betatransfem_share, that is the beta transformation of the fem_share dependent variable, as described in section 4.1.2.

Two other results are included in table 11. On the one hand, the share of STEM graduates among the total population of graduates correlates negatively with the patent gender gap, i.e. the more STEM graduates there are the more severe the patent gender gap is. Indeed, estimates are negative and significant at the 5% level in all specifications that include this variable as an independent variable – specifications (1), (2) and (7). The estimated effect of this variable on the patent gender gap varies between $e^{-0.026} = 0.9743$, that is a 2.57% decrease – specification (2) – and $e^{-0.039} = 0.9618$, that is a 3.82% decrease in the ratio of the proportion of women patentees to men patentees. On the other hand, the share of female STEM graduates among the total STEM graduates – or, put differently, the gender gap in STEM graduates – correlates positively with the proportion of women among the total patenting population. The estimated coefficient on the variable F_Share_STEM is both positive and significant at the 5% level in all specifications that include this variable – specifications (1), (2) and (7). The estimated effect of this variable on the patent gender gap varies between $e^{0.019} = 1.0192$, that is a 1.92% increase – specification (7) – and $e^{0.023} = 1.0233$, that is a 2.33% positive change in the patent gender gap. Hence, it appears that narrowing the gender gap in STEM academic disciplines reduces the patent gender gap. Namely, the higher the share of women in STEM-related academic fields, the higher the proportion or share of women in the patenting population, which validates the third hypothesis. These results indicate that encouraging more women towards a higher level of education without increasing their participation to STEM academic fields does not improve the patent gender gap in patenting, and a combination of both increasing a higher educational achievement and encouraging women to undertake a STEM education is necessary.

However, the effect of implementing such incentives for women in STEM fields is uncertain, as the educational gender gap is not the only possible explanation for the underrepresentation of women in patenting. The access to STEM-related jobs also suffers from a blatant gender gap, and the access to this segment of the job market is biased in men's favour. Based on American data, Beede et al. (2011) assert that women with a STEM degree are also less likely than men to work in a STEM occupation, and that women are likely to work in healthcare and education. Hence, there is also a STEM-employment gender gap, for which Beedle et al. provide several explanations: the absence of female model – itself due to the underrepresentation of women in STEM employment, creating a vicious circle – gender-based stereotypes, impaired flexibility due to family obligations, corroborating the observations made by Wynarczyk & Renner (2006) and Weinkopf (2014). Gender stereotypes and the lack of early exposure to innovation-oriented programmes also seem to hinder women's patenting (Couch, Estabrooks, & Skukauskaite, 2018).

Furthermore, the estimated coefficient for the GII variable in table 11 is highly negative as well as statistically significant at the 5% level in all specifications (5), (6) and (7). As mentioned previously, the GII is calculated in such a way that the higher the index, the more blatant the gender inequalities are. Hence, the negative and significant estimated coefficient points towards the validation of the fourth hypothesis.

6. Conclusion

In this paper, the geographical patterns of the patenting gender gap as well as the educational driving forces of patenting have been empirically investigated. The sample used in these investigations comprises 27 European countries belonging either to the European Union or the OECD, or both, which are divided into 267 NUTS2 statistical regions. From this sample, the map visualisation of the data – the share of women among the regional population of patent applicants – shows the existence of a city effect that narrows the patent gender gap. However, this effect does not seem to operate fully in central European countries, due to employment-related gender gaps affecting women's productivity and work-life balance. In addition, the relatively higher participation of women in patenting in European regions formerly belonging to the Soviet bloc as well as the existence of a handful of cross-border spillovers between geographical clusters indicate that culture and history are other factors affecting the patent gender gap on the European continent. Complementing the analysis of the geographical patterns, the development and application of beta-regression and panel models show that one of the key elements for narrowing the patent gender gap is to reduce the educational gender gap at the highest education level in STEM fields, that is to increase the participation of women to scientific fields. However, narrowing the education gender gap in STEM is unlikely to be the one-size-fit-all solution in achieving gender parity in patenting on the European continent, as other gender gaps – related to employment, stereotypes and the role of women with regard to their family – would remain unsolved by merely encouraging more women to pursue STEM education. Hence, the most appropriate policies in narrowing the patent gender gap are policies that take a holistic approach of gender gaps rather than focusing on a gender gap in particular field.

Takeaways of this study can be found both for public policy makers and private businesses. The first key takeaway of this study for policy makers is that incentive schemes encouraging innovation and patenting by women should be emphasised in large cities and their surroundings, as they seem to positively influence the frequency of women patenting. As a result, the efficiency of women patenting policies could be higher compared to less urbanised areas. However, some important deviations from the city effect exist, in particular in countries where women suffer from larger inequalities when it comes to the access to the labour market by taking part-time positions much more often than men, which means that these countries should focus even more on devising and implementing policies enabling women to take on or keep full-time positions regardless of their family situations and parenthood. In addition, policy makers should not only pursue a narrower gender gap in education in general but also implement incentives aiming at increasing the share of women in STEM disciplines, the educational gender gap is already quite narrow in many non-STEM educational fields and women are still widely underrepresented in STEM disciplines. This would result in improved career-progression potential as well as an increased scientific output production by women (Bertocchi & Bozzano, 2019). The mere pursuit of more women in tertiary education as a whole is likely to miss the target while letting the gender gap in STEM narrow further. Lastly, gender equality in its various dimensions – access to the labour market, working terms and conditions, family duties, etc. – are an overarching goal that has to be pursued in order to improve the patenting output of the feminine half of the European population. When it comes to businesses, a possible takeaway from this study is that private businesses should encourage more diverse teams, that is the inclusion of more women in their scientific teams, as diversity promotes equality, which in turns fosters innovation and eventually can result in a higher innovation and patenting output. Another takeaway for businesses is the recommended implementation of schemes providing support to women at the workplace, for instance by offering flexible working hours and/or environments better enabling offering women a better work-life balance to mitigate the dilemma between family obligations and work duties working mothers can be faced with, and by offering continuous training programs aiming at narrowing the human capital gaps between men and women especially in technical and scientific industries and business units.

Contrasting with this study, researchers have also argued that considering the mere number or frequency of patenting by women does not bring enough nuancing to the analysis of gender disparities in patenting. For instance, Sugimoto et. al (2015) focused on the organisational context in which women patenting occurred as well as the gender disparities in the impact of patenting. They found that women patenting is less likely to occur in corporate and governmental environments but more likely in academic institutions, and that the technological impact of women's patents is comparatively lower than men's patenting. Consequently, Hence, further studies on gender disparities in light of patent quality and productivity can only be encouraged. Lastly, the suitability of patent data as a measure of innovation is largely discussed, as not all innovations are patented (see for example Griliches, 1998) and patents are not a necessary condition to show the existence of innovation (Morand & Manceau, 2009).

References

Akinyemi, J. O. (2016). Creating a Virtuous Circle: Developing Intellectual Property and Innovation Policy and Strategies in Public Research and Development Institutions and Universities to Facilitate the Transfer of Technology to Industry. Retrieved November 30, 2020, from https://www.wipo.int/edocs/mdocs/africa/en/ompi_inn_cas_16/ompi_inn_cas_16_t_1.pdf

Beede, D. N., Julian, T. A., Langdon, D., McKittrick, G., Khan, B., & Doms, M. E. (2011). Women in STEM: A Gender Gap to Innovation. *SSRN Electronic Journal*. doi:10.2139/ssrn.1964782

Bertay, A. C., Dordevic, L., & Sever, C. (2020). Gender Inequality and Economic Growth. *IMF Working Papers, 20*(119). doi:10.5089/9781513546278.001

Bertocchi, G., & Bozzano, M. (2019). Gender Gaps in Education. Retrieved October 10, 2020, from http://ftp.iza.org/dp12724.pdf

Bettencourt, L. M., Lobo, J., Helbing, D., Künhert, C., & West, G.B. (2007). Growth, Innovation, Scaling and the Pace of Life in Cities. Proceedings of the national academy of sciences, 104(17), 7301-7306. doi:10.1073/pnas.0610172104.

Buis, M. L. (2010). Analyzing Proportions [Power Point Slides]. Retrieved October 18, 2020, from http://fmwww.bc.edu/repec/dsug2010/buis.pdf

Cameron, A. C., & Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources, 50*(2), 317-372. doi:10.3368/jhr.50.2.317

Carlino, G. A., Chatterjee, S., & Hunt, R. M. (2007). Urban density and the rate of invention. *Journal of Urban Economics, 61*(3), 389-419. doi:10.1016/j.jue.2006.08.003

Cornell University, INSEAD, and WIPO (2017): *The Global Innovation Index 2017: Innovation Feeding the World*, Ithaca, Fontainebleau, and Geneva. ISSN 2263-3693. ISBN 979-10-95870-04-3

Couch, S., Estabrooks, L. B., & Skukauskaite, A. (2018). Addressing the Gender Gap among Patent Holders through Invention Education Policies. *Technology & Innovation, 19*(4), 735-749. doi:10.21300/19.4.2018.735

Durà, A., Camonita, F., Berzi, M., & Noferini, A. (2018). Euroregions, Excellence and Innovation across EU borders: A Catalogue of Good Practices. Retrieved October 29, 2020, from https://ec.europa.eu/futurium/en/system/files/ged/recot_crii_catalogue_0.pdf

Edgeworth, F. Y., & Marshall, A. (1891). Principles of Economics. *The Economic Journal, 1*(3), 611. doi:10.2307/2956141

European Central Bank. (2017). How does innovation lead to growth? Retrieved November 15, 2020, from https://www.ecb.europa.eu/explainers/tell-me-more/html/growth.en.html

European Commission. Directorate General for Justice. (2013). Barcelona objectives :the development of childcare facilities for young children in Europe with a view to sustainable and inclusive growth. Publications Office. https://doi.org/10.2838/43161

Ferrari, S., & Cribari-Neto, F. (2004). Beta Regression for Modelling Rates and Proportions. *Journal of Applied Statistics, 31*(7), 799-815. doi:10.1080/0266476042000214501

Florida, R., Adler, P., & Mellander, C. (2016). The city as innovation machine. *Regional Studies, 51*(1), 86-96. doi:10.1080/00343404.2016.1255324

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in Cities. *Journal of Political Economy, 100*(6), 1126-1152. doi:10.1086/261856

Godin, B. (2019). *The invention of technological innovation: Languages, discourses and ideology in historical perspective*. Cheltenham, UK: Edward Elgar Publishing.

Griliches, Z. (1998). Patent Statistics as Economic Indicators: A Survey. Retrieved November 15, 2020, from https://www.nber.org/books-and-chapters/rd-and-productivity-econometric-evidence/patent-statistics-economic-indicators-survey

Gui-Diby, S. L., Pasali, S. S., & Rodriguez-Wong, D. (2017). *What's gender got to do with firm productivity? Evidence from firm level data in Asia* (Working paper No. WP/17/01). Bangkok: Macroeconomic Policy and Financing for Development Division (MPFD). Retrieved November 14, 2020, from https://www.unescap.org/sites/default/files/publications/MPFD_WP-17-01_Gender%20and%20Productivity.pdf

Hunt, V., Yee, L., Prince, S., & Dixon-Fyle, S. (2020). Delivering through diversity. Retrieved November 15, 2020, from https://www.mckinsey.com/business-functions/organization/our-insights/delivering-through-diversity?cid=other-eml-nsl-mip-mck-oth-1802

Innovation and Intellectual Property. (n.d.). Retrieved November 15, 2020, from https://www.wipo.int/ip-outreach/en/ipday/2017/innovation_and_intellectual_property.html

Innovation, Creativity and the Gender Gap. (n.d.). Retrieved November 10, 2020, from https://www.wipo.int/ip-outreach/en/ipday/2018/innovation_creativity_gender_gap.html

Jaffe, A. B., Trajtenberg, M., & Fogarty, M. S. (2000). Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors. *American Economic Review, 90*(2), 215-218. doi:10.1257/aer.90.2.215

Lax Martínez, G., Raffo, J., & Saito, K. (2016). Identifying the Gender of PCT Inventors. Retrieved October 16, 2020, from https://www.wipo.int/publications/en/details.jsp?id=4125

Liberda, B., & Zajkowska, O. (2017). Innovation and Entrepreneurship Policies and Gender Equity. *International Journal of Contemporary Management,* (1). doi:10.4467/24498939ijcm.17.002.7046

Makkonen, T., & Inkinen, T. (2013). Innovative Capacity, Educational Attainment and Economic Development in the European Union: Causal Relations and Geographical Variations. *European Planning Studies, 21*(12), 1958-1976. doi:10.1080/09654313.2012.722968

Maraut, S., Dernis, H., Webb, C., Spiezia, V., & Guellec, D. (2008). The OECD REGPAT Database: A Presentation. Retrieved October 16, 2020, from https://dx.doi.org/10.1787/241437144144

Matteazzi, E., Pailhé, A., & Solaz, A. (2017). Part-time employment, the gender wage gap and the role of wage-setting institutions: Evidence from 11 European countries. *European Journal of Industrial Relations, 24*(3), 221-241. doi:10.1177/0959680117738857

McManus, P. A. (2011). Introduction to Regression Models for Panel Data Analysis [Power Point Slides]. Retrieved October 24, 2020, from https://ssrc.indiana.edu/doc/wimdocs/2011-10-07_mcmanus_panel_slides.pdf

Miller, S. (2019). Unequal Career Advancement Fuels Gender Pay Gap. Retrieved November 30, 2020, from https://www.shrm.org/ResourcesAndTools/hr-topics/compensation/Pages/unequal-career-advancement-fuels-gender-pay-gap.aspx

Morand, P., & Manceau, D. (2009). Pour une nouvelle vision de l'innovation. Rapport officiel, Paris, La Documentation française. Retrieved November 12, 2020, from http://www.eurosfaire.prd.fr/7pc/doc/1244626060_autre_vision_innovation_escp_europe_avril_2009.pdf

OECD. (2004). Patents and innovation: trends and policy challenges. Retrieved November 15, 2020, from http://www.oecd.org/science/inno/24508541.pdf

Peterlini, O. (2007). The South Tyrol Autonomy in Italy, International Conference "One country, two systems, three legal orders - perspectives of evolution". Retrieved October 29, 2020, from https://www.researchgate.net/publication/316861601_The_South_Tyrol_Autonomy_in_Italy_International_Conference_One_country_two_systems_three_legal_orders

Phillips, K. W. (2014). How Diversity Makes Us Smarter. Retrieved November 15, 2020, from https://www.scientificamerican.com/article/how-diversity-makes-us-smarter/

Pratt, A. C. (2008). Creative cities: The cultural industries and the creative class. *Geografiska Annaler: Series B, Human Geography, 90*(2), 107-117. doi:10.1111/j.1468-0467.2008.00281.x

Schumpeter, J. A. (1942). *Capitalism, Socialism and Democracy*. London: Taylor and Francis.

Scott, A. J. (2006). Entrepreneurship, Innovation and Industrial Development: Geography and the Creative Field Revisited. *Small Business Economics, 26*(1), 1-24. doi:10.1007/s11187-004-6493-9

Shook, E., & Sweet, J. (2019). Equality = Innovation. Getting to Equal 2019: Creating a culture that drives innovation. Retrieved November 30, 2020, from https://www.accenture.com/_acnmedia/Thought-Leadership-Assets/PDF/Accenture-Equality-Equals-Innovation-Gender-Equality-Research-Report-IWD-2019.pdf

Smithson, M., & Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. Psychological Methods, 11(1), 54–71. https://doi.org/10.1037/1082-989x.11.1.54

Sohn, C., Reitel, B., & Walther, O. (2009). Cross-Border Metropolitan Integration in Europe: The Case of Luxembourg, Basel, and Geneva. doi:10.1068/c0893r

Stephan, P. E., & Elganainy, A. A. (2006). The Entrepreneurial Puzzle: Explaining the Gender Gap. *SSRN Electronic Journal*. doi:10.2139/ssrn.975953

Stock, J. H., & Watson, M. W. (2020). *Introduction to econometrics* (4th ed.). Harlow, England: Pearson. ISBN 9781292264523

Sugimoto, C. R., Ni, C., West, J. D., & Larivière, V. (2015). The Academic Advantage: Gender Disparities in Patenting. *Plos One, 10*(5). doi:10.1371/journal.pone.0128000

Swearingen, C. J., Melguizo Castro, M. S., & Bursac, Z. (2011). Modeling Percentage Outcomes: The %Beta_Regression Macro. Retrieved October 21, 2020, from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.353.3735&rep=rep1&type=pdf

Torres-Reyna, O. (2007). Panel Data Analysis Fixed and Random Effects using Stata. Retrieved October 25, 2020, from https://www.princeton.edu/~otorres/Panel101.pdf

UNESCO science report: Towards 2030. (2016). Paris: UNESCO.

Weinkopf, C. (2014). Women's Employment in Germany. *Revue De L'OFCE, 2*, 133-189. doi:10.3917/reof.133.0189

WIPO IP Facts and Figures 2019. (2019). Retrieved November 3, 2020, from https://www.wipo.int/edocs/pubdocs/en/wipo_pub_943_2019.pdf. ISBN: 978-92-805-3129-9

Wynarczyk, P., & Renner, C. (2006). The “gender gap” in the scientific labour market: The case of science, engineering and technology-based SMEs in the UK. Equal Opportunities International, 25 (8), 660-673. https://doi.org/10.1108/02610150610719128

Appendices

Appendix A – Calculation of Beta-regression parameters

Using the mean R function, the mean of fem_share is: $E(y) = 0.09951998 = \mu$

Using the variance R function, the variance of fem_share is: $V(y) = 0.007983993$

Hence,

$$\frac{V(y)}{\mu(1-\mu)} = \frac{1}{1+\phi}$$

$$1+\phi = \frac{\mu(1-\mu)}{V(y)}$$

$$\phi = \frac{\mu(1-\mu)}{V(y)} - 1$$

$$\phi = \frac{0.09951998(1-0.9951998)}{0.007983993} = 10.22442787$$

Given that (see Ferrari & Cribari-Neto, 2004):

$$\alpha = \left[\frac{1-\mu}{V(y)} - \frac{1}{\mu}\right]\mu^2$$

and

$$\beta = \left[\frac{1-\mu}{V(y)} - \frac{1}{\mu}\right]\mu(1-\mu)$$

It follows that:

$$\alpha = \left(\frac{1-0.09951998}{0.007983993} - \frac{1}{0.09951998}\right)0.09951998^2 = 1.01753$$

and

$$\beta = \left(\frac{1-0.09951998}{0.007983993} - \frac{1}{0.09951998}\right)0.09951998(1-0.09951998) = 9.20689$$

Appendix B – Breusch-Pagan test for heteroskedasticity control

Assuming the following specification:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_i x_i + \varepsilon_i$$

- Null hypothesis ($H0$): homoskedasticity, i.e.

$$H0: V(\varepsilon_i | x_i) = \sigma^2$$

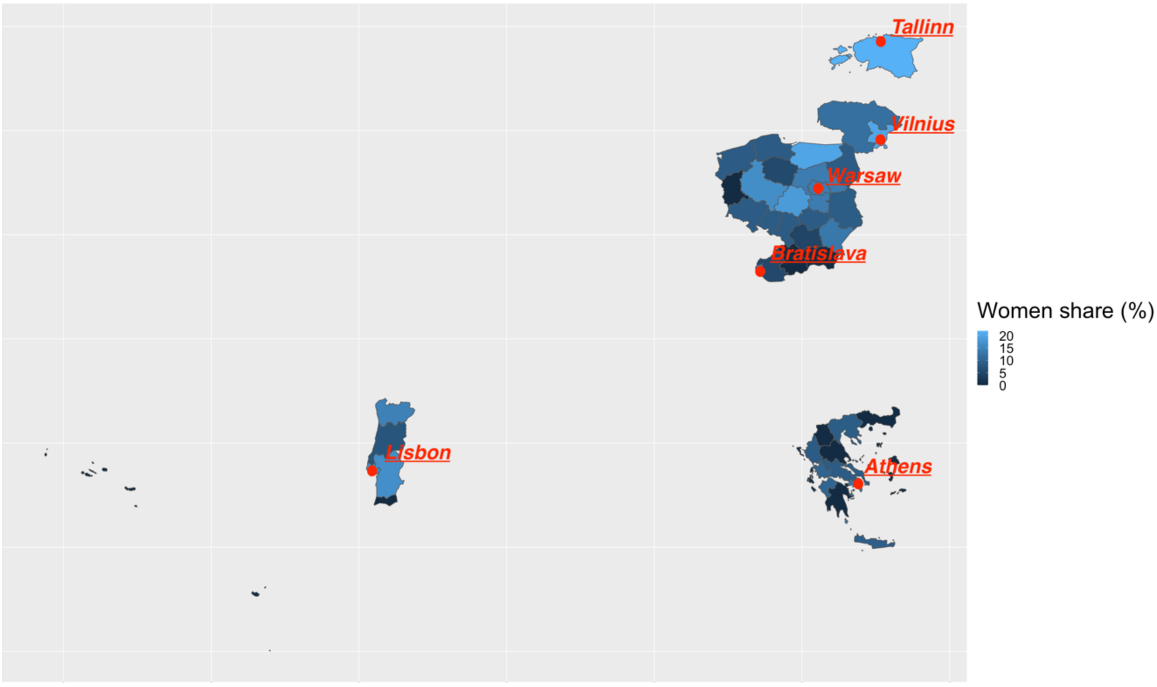
- Estimation of the squared of residuals

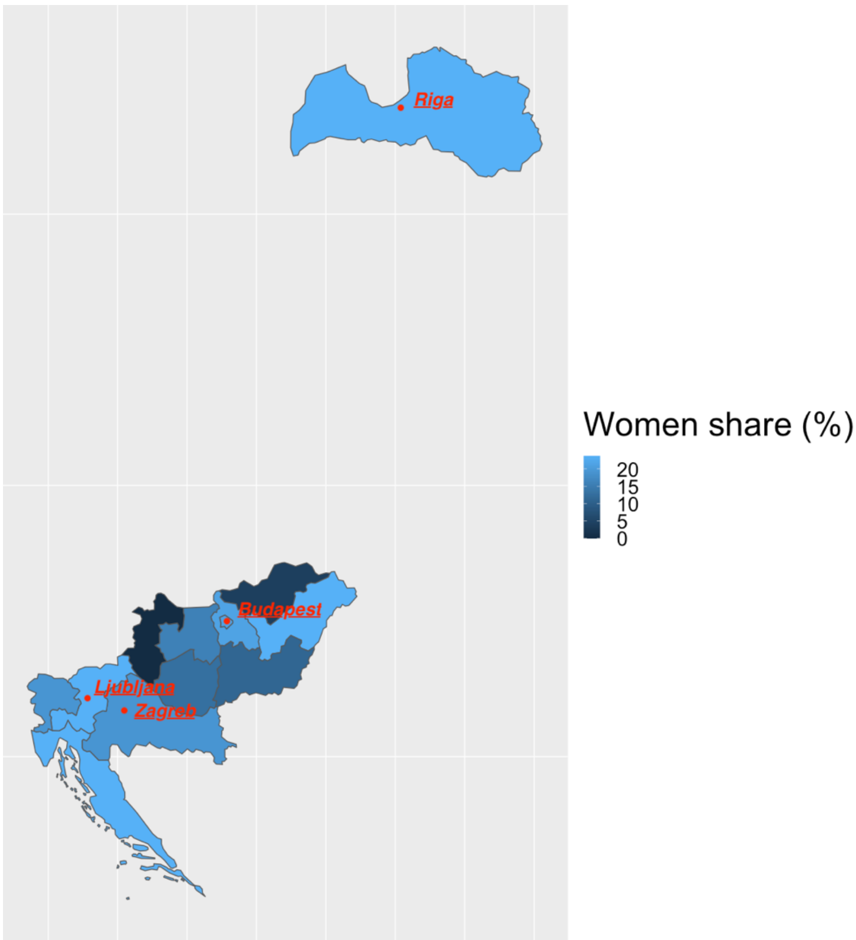
$$\widehat{u_i}^2 = \gamma_0 + \gamma_1 x_1 + \cdots + \gamma_i x_i + e => R^2_{\widehat{u_i}^2}$$

- Test statistic

$$LM = nR^2_{\widehat{u_i}^2} \sim \chi^2_i$$

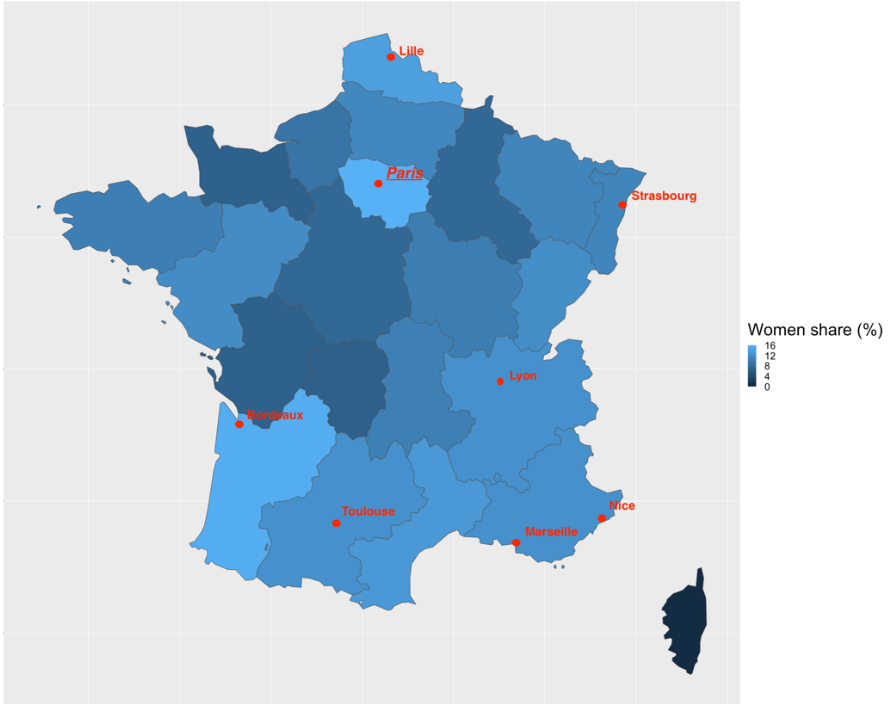
Appendix C – Women share within the southern and eastern as well as the eastern European clusters



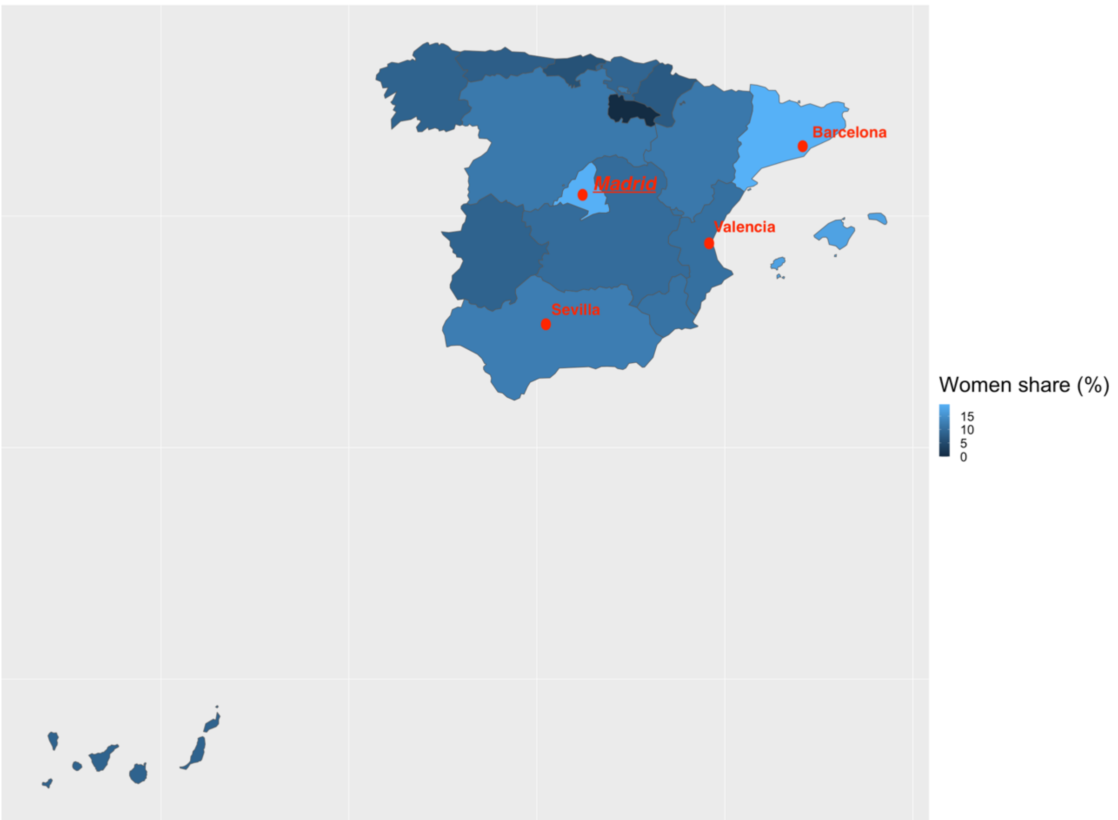


Appendix D – Maps of all countries

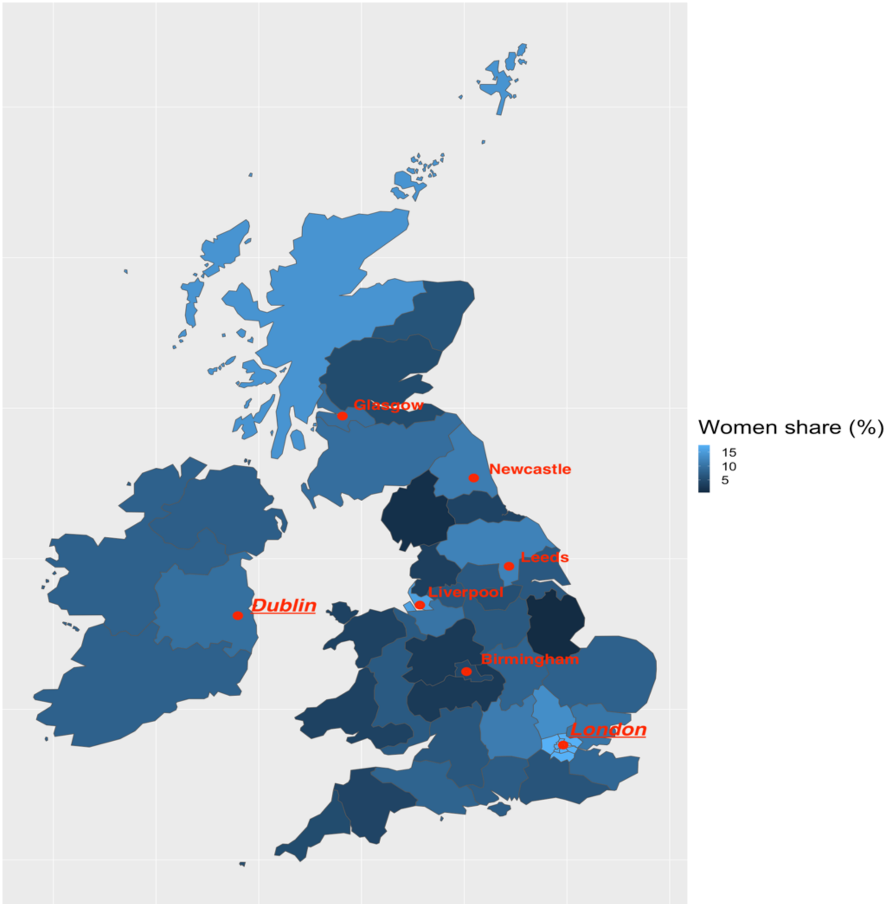
France



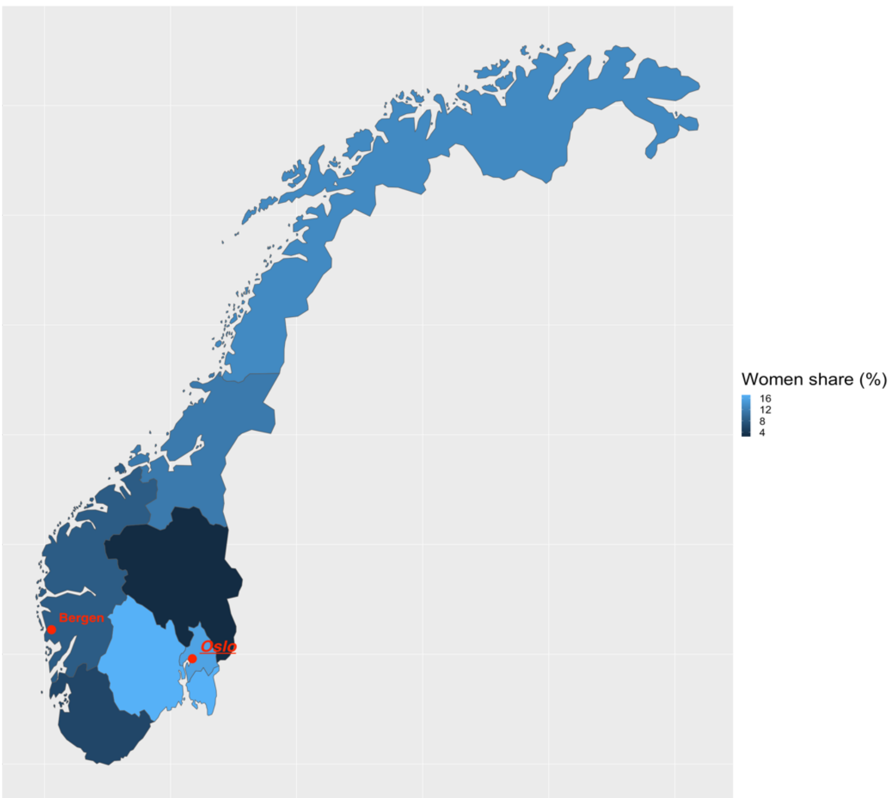
Spain



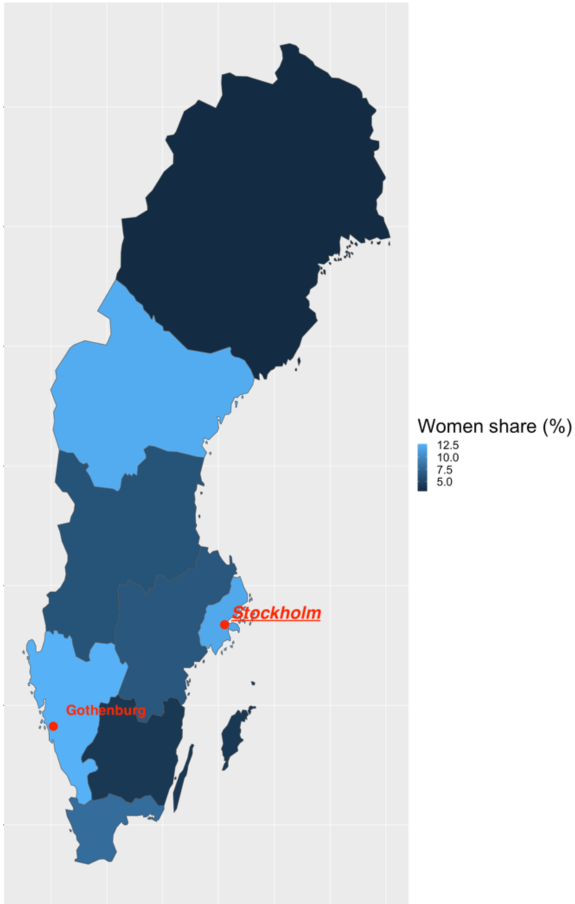
United Kingdom & Ireland



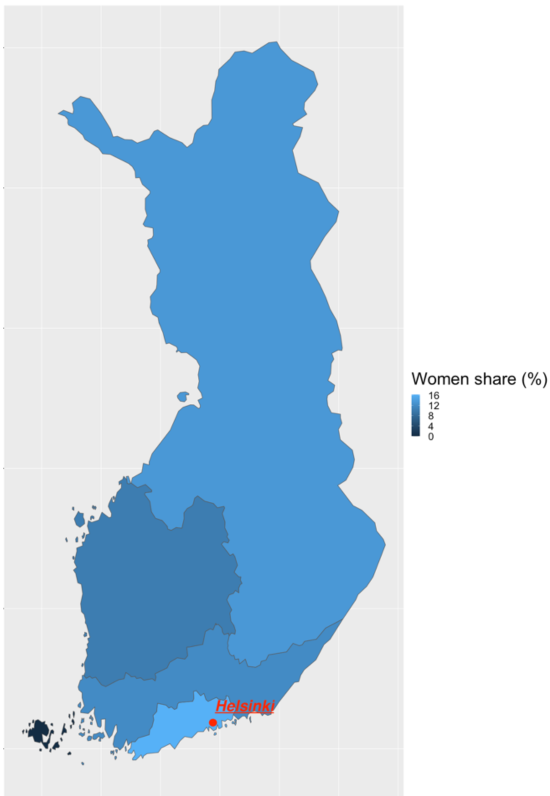
Norway



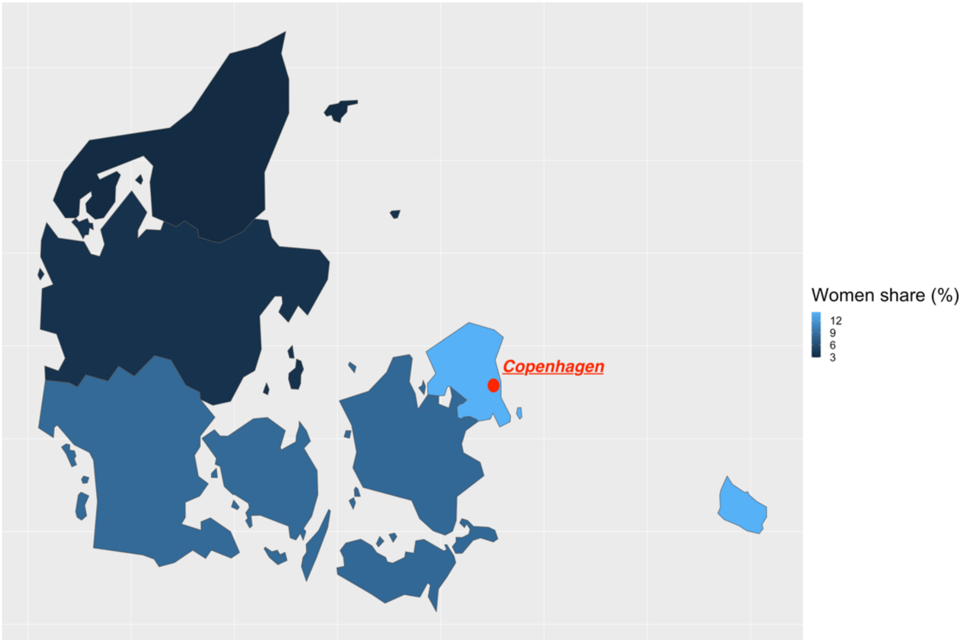
Sweden



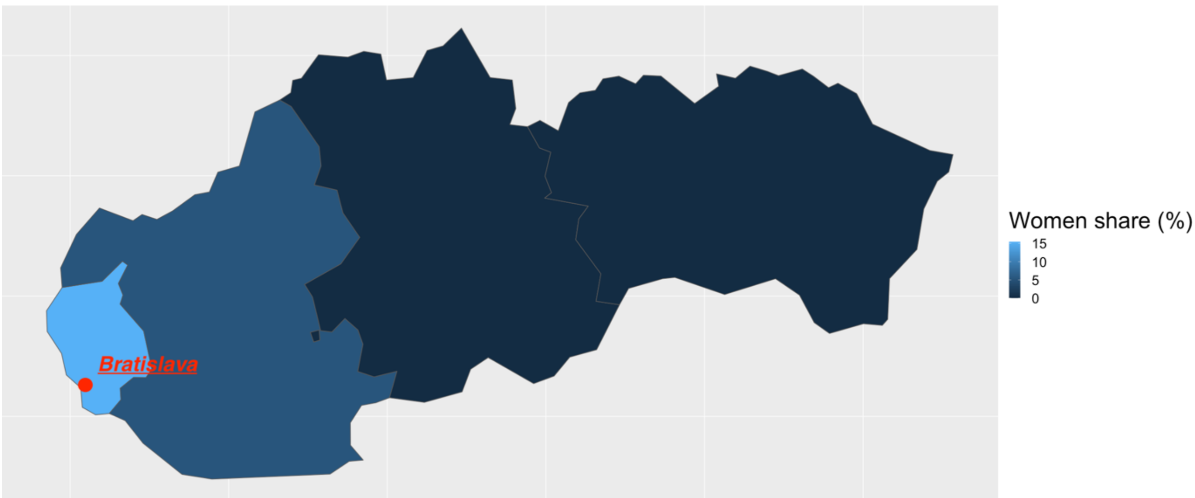
Finland



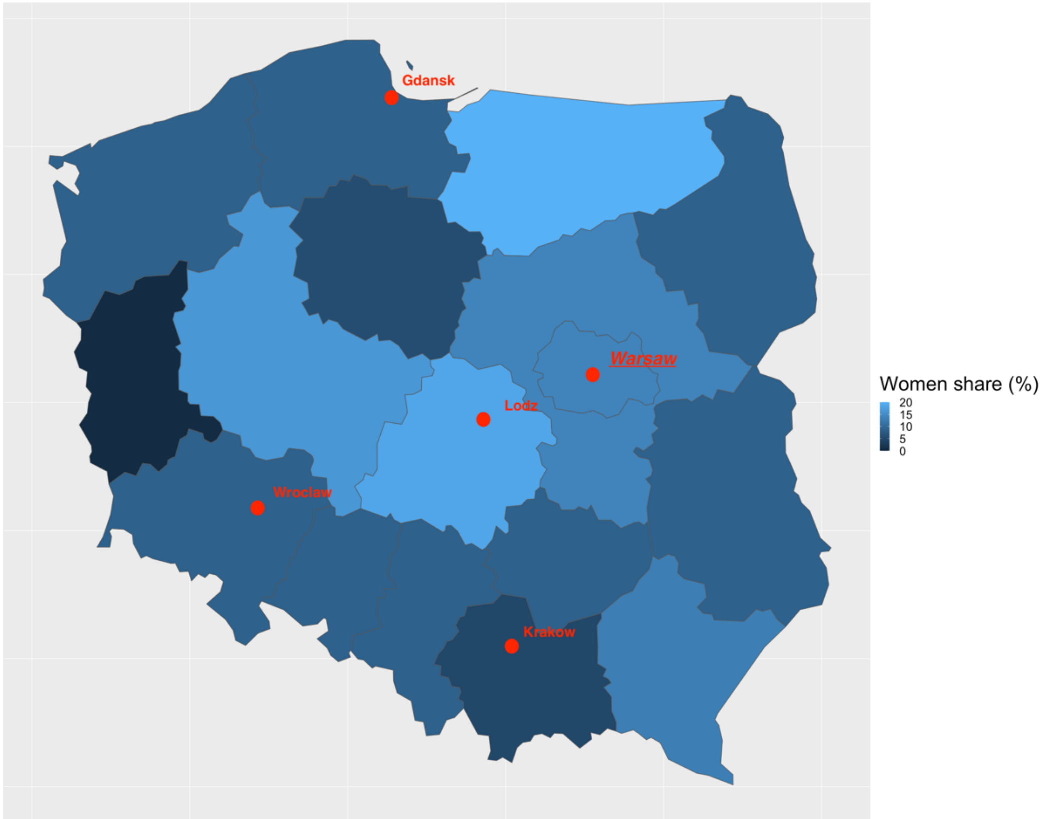
Denmark



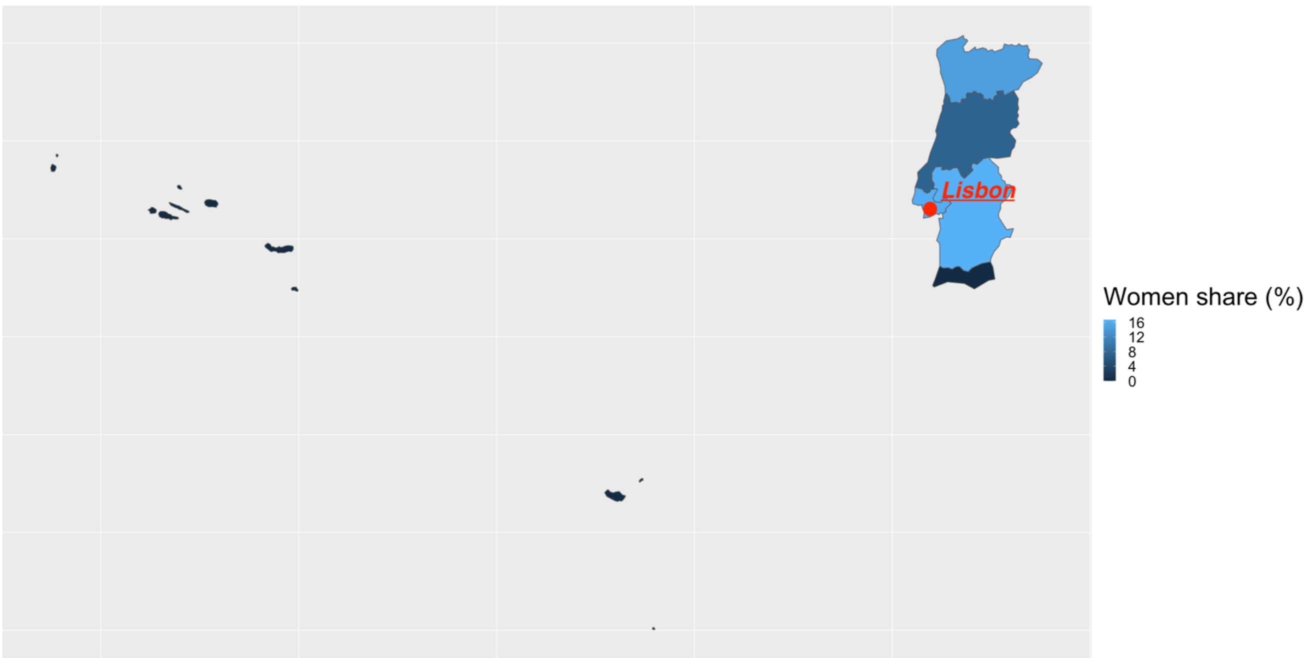
Slovakia



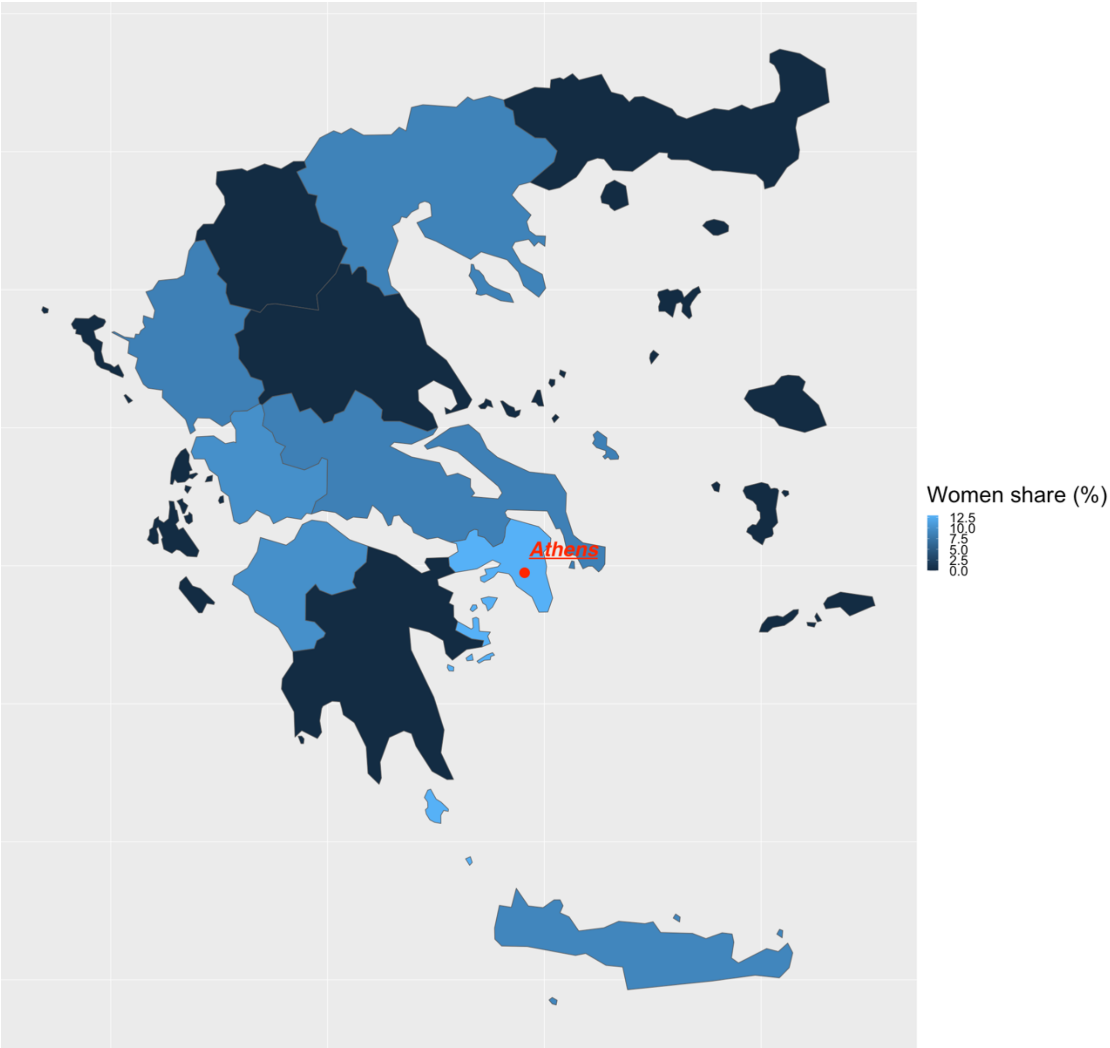
Poland



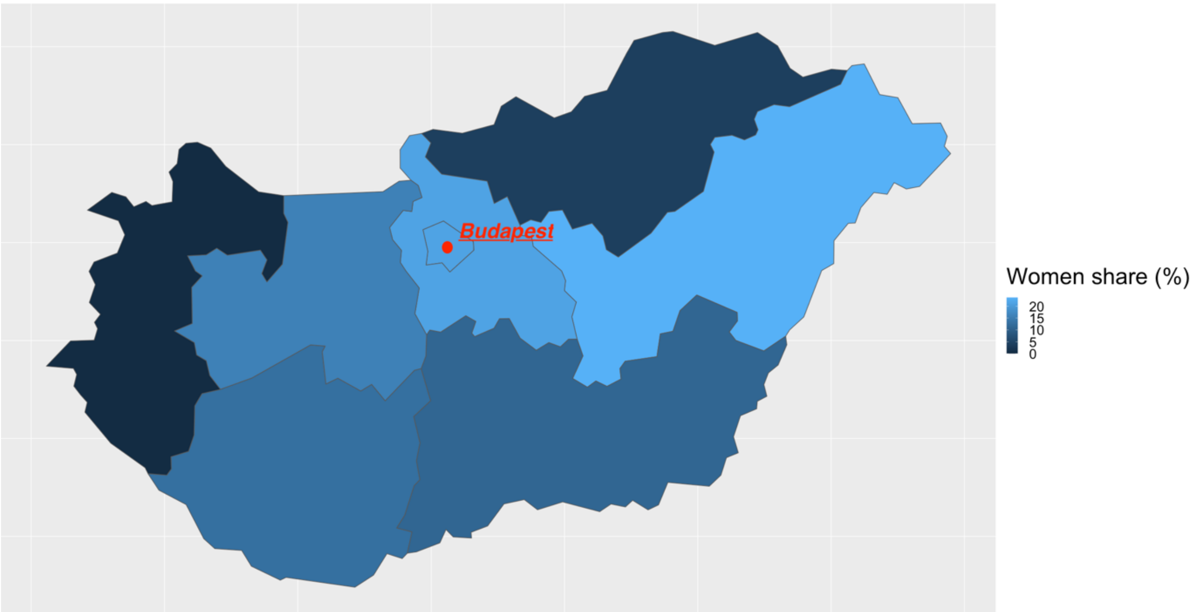
Portugal



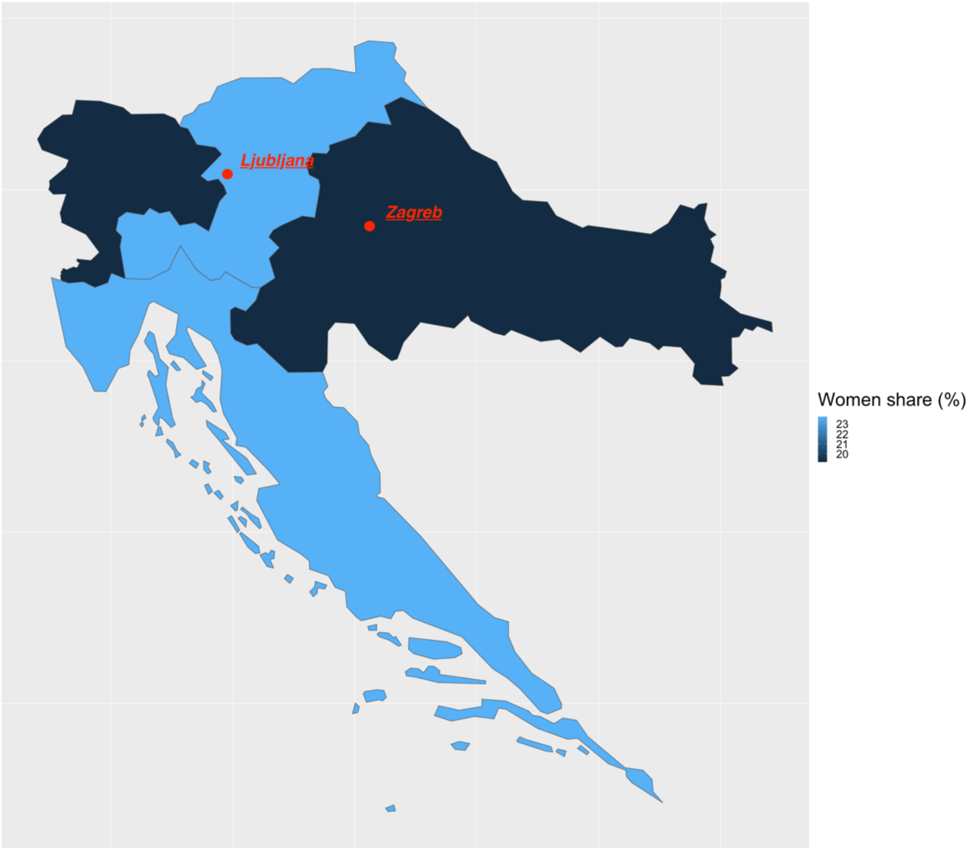
Greece



Hungary



Croatia & Slovenia



The Baltics