NORTHWESTERN UNIVERSITY

Essays on Innovation Policy

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Managerial Economics and Strategy

By

Rainer Widmann

EVANSTON, ILLINOIS

September 2018

Abstract

This dissertation focuses on policy issues in the area of Economics of Innovation. As developed countries are becoming increasingly reliant on innovation for economic growth, it is important to enhance our understanding of how public policies affect innovation.

The first chapter examines the effect of government research grants on firms' patenting outcomes. Discontinuities in the funding decisions of the Austrian Research Promotion Agency (FFG) allow me to study the effect of public funding in a large sample of Austrian firms. My estimates suggest that a government research grant increases the propensity to file a patent application with the European Patent Office within 4 years by around 10 percentage points. Stronger effects appear for established firms of advanced age. I present evidence that established firms undertake ambitious research projects when they receive grants. Finally, I interpret the findings in an "exploration vs. exploitation" model in which the government agency addresses inefficiency in the direction of research.

In the the second chapter, I study the relationship between personal income tax rates and the residential location choice of inventors in Switzerland. Exploiting sharp differences in tax rates for top-income earners across state borders, I find an elasticity of the number of inventors in a municipality with respect to the net-of-tax rate (after-tax income) of around 4.6. This estimate is considerably higher than the elasticities found in previous studies of inventor mobility. Tax policies at the local level, where inventors may lower their tax burden by relocating over short distances, may have particularly strong effects. In addition, I study how location choices depend on other non-pecuniary local amenities. Finally, I document that inventors who have a longer commute are more likely to transfer to a workplace closer to their residence over time. Hence, using

income tax policies to draw the residence of inventors closer may have positive spillover effects.

The third chapter studies the allocation of R&D subsidies under different modes of trade integration. Two governments pay subsidies to domestic firms that are either technology leaders or technology followers. The model predicts that trade integration leads to increases in aggregate research spending and the share of business R&D funded by government. While subsidies are directed at technology leader firms when markets are separated, trade integration leads to a more even distribution of subsidies that benefits technology follower firms. I show that these findings are broadly consistent with international policy trends. In particular, over the past decades, indiscriminate R&D tax credits have increased in importance relative to direct, discretionary funding of business R&D.

Contents

List of Tables

List of Figures

1 The effect of government research grants on firm innovation: theory and evidence from Austria

1.1 Introduction

Since firms cannot appropriate all gains generated by inventions, private investment in Research & Development falls short of the socially optimal level of investment (see, for example, Bloom et al. 2013, Hall 2002, Jones and Williams 1998, Jaffe 1986, Nelson 1959). For this reason, all developed countries have put policies in place that support and enable firm innovation. Aside from maintaining the patent system, governments regularly subsidize R&D in order to raise the rate of innovation. Subsidies are provided directly through grants, allocated at the discretion of government agencies, and indirectly through tax credits. In this chapter, I study the effects of a particular directed government research grant program.

Despite their widespread use, evidence on the effectiveness of directed subsidies is inconclusive (see Zúñiga-Vicente et al. 2014, discussed in section 1.2), or supports the view that positive effects may be confined to younger or smaller firms (see Howell 2016, Bronzini and Iachini 2014, discussed in section 1.2). Critics contend that government research grant programs are prone to failure for several reasons: first, government agencies may be biased to award grants to "winners" whose projects would have been undertaken and turned into a success even without a grant. Second, critics question the ability of bureaucrats to evaluate industrial research projects. Third, government agencies that enjoy discretion may be plagued by capture by the firms they are destined to serve. Notwithstanding these arguments, directed grants may offer distinctive advantages over undirected tax credits. Funding can be targeted at research projects at the margin, thereby potentially stimulate innovation at lower cost. Furthermore, funding can be targeted at ambitious research projects that have high social externalities.

The study in this chapter is based on 2619 applications to the Austrian Research Promotion Agency (FFG) across 1936 firms between 2002 and 2005. To identify the effect of the research grant, I exploit a steep, almost discontinuous increase in the dependence of funding approval on the agency's internal evaluation score. I find that FFG research grants have an effect of around 10 percentage points on the propensity to file a patent application with the European Patent Office within 4 years. The effect is stronger for firms above the median age in the sample, which is 5 years. I present additional survey evidence, collected during a review of the FFG by a team of innovation policy experts in 2003, that suggests that established firms increase risk and ambition of research projects when they receive a grant. Consistent with this evidence, I find that for firms with prior patent filings, funding approval is correlated with heavier utilization of technological knowledge novel to the firm after the grant application. Furthermore, I show that, in spite of the fact that the agency only rejects the bottom third of applications, marginal patents produced due to research grants appear non-trivial with regard to conventionality and number of citations received for firms older than 5 years.

I interpret the evidence in an "exploration vs exploitation" model, where established firms face the choice between exploring novel research lines and exploiting old research lines. After an initial "breakthrough", research lines spawn a series of incremental projects of declining value. Assuming that the social value of a research project scales linearly to the private value, I show that firms over-exploit compared to the social optimum. The social planner can implement the optimal research policy by offering research grants for explorative projects. Young firms have no old research lines to exploit and are therefore not subject to this mechanism. In addition, I use the model to help elucidate the relative effectiveness of research grants and undirected tax credits in this setting. The model relates to recent papers on inefficieny in the direction of research (see Bryan and Lemus 2016, Hopenhayn and Squintani 2016) and to the literature on the role and optimal design of innovation policy (see Akcigit et al. 2013, Acemoglu et al. 2014, Besanko and Wu 2010).

The rest of the chapter is organized as follows. In section 1.2, I discuss related literature. Section 1.3 describes the Austrian Research Promotion Agency FFG and the data. In Section 1.4, I lay out the empirical strategy. Section 1.5 contains the results of the effect of research grants. Section 1.6 presents the model. Section 1.7 concludes.

1.2 Related literature

This chapter contributes to several strands of the literature. First, it contributes to the literature evaluating directed R&D subsidy programs. Howell (2016), Bronzini and Iachini (2014), Einiö (2014) and Bronzini and Piselli (2016) were the first studies to present quasi-experimental evidence on this subject. Since selection for funding in R&D subsidy programs depends on the quality of the submitted project, selection bias is a serious concern in this context. Howell (2016) uses a Regression Discontinuity Design to examine the US SBIR program and finds a positive effect of Phase-1 grants on quality-adjusted patent outcomes[1], which is stronger for firms that are first-time applicants to the program. She corroborates this finding by showing that young recipient firms are subsequently more likely to receive outside venture capital. Howell (2016) argues that the provision of funds to young companies helps alleviate financial frictions present in the market for venture capital. Bronzini and Iachini (2014), also in a RD design, study a North-Italian research grant program and find a positive effect on R&D expenditures for small firms. In a recent follow-up paper,

[1] Patents are weighted by citation counts.

Bronzini and Piselli (2016) find a positive effect on the number of patent applications for the same program, with stronger effects for small firms. They only find an effect on the propensity to file a patent application for small firms. They also attribute their results to financial frictions faced by small firms. Utilizing exogenous variation in subsidies generated by population-density rules, Einiö (2014) finds evidence of a positive effect of R&D subsidies on R&D expenditures, employment, sales and productivity in a large sample of Finish firms. There is a sizeable number of descriptive studies that have examined the effects of different R&D subsidy programs on R&D expenditures, with mixed results. Zúñiga-Vicente et al. (2014) survey 76 studies and conclude that evidence on the effect of direct subsidies on R&D expenditures is inconclusive. The authors suppose that differences in the designs of the subsidy programs and differences in the studied firm populations account for the heterogeneity in the findings.[2] While the effects of direct subsidies on R&D expenditures have been studied widely, there are only a handful of studies on the effect of direct subsidies on other measures of innovative activity. Besides the studies by Howell (2016), Bronzini and Piselli (2016) and Einiö (2014) mentioned earlier, I am only aware of studies by Lerner (1999), who finds that US-SBIR awardees grew faster than their matched peers, and Bérubé and Mohnen (2009), who find a positive effect on the introduction of new products in a matched sample of Canadian firms.[3] A related paper is Branstetter and Sakakibara (2002), who investigate the effect of participation in Japanese R&D consortia.

The findings in this chapter complement results from the literature on R&D tax credits. There is an extensive literature assessing the impact of R&D tax credits on R&D expen-

[2]Important studies included in the survey are Wallsten (2000) and Lach (2002) who do not find an effect of subsidies on R&D expenditures, and Lichtenberg (1988), Almus and Czarnitzki (2003), González et al. (2005), Hussinger (2008) who find positive and statistically significant effects. Takalo, Tanayama and Toivanen (2013) study R&D subsidies in a structural model of the Finish agency Tekes.

[3]Bronzini and Piselli (2016) remark in their review of the literature that it is puzzling that so few studies of the effect on innovation outputs exist.

ditures, both on the macroeconomic level and on the firm level (surveyed in Becker 2014). Most notably, Dechezleprêtre et al. (2016) employ a combined RD/Difference-in-Difference strategy to investigate the effect of a R&D tax credit on small British firms around an eligibility-threshold. They find positive effects on R&D expenditures and patenting activity. Other studies that have examined the effect of R&D tax credits on innovation output are Capellen et al. (2012), who do not find an effect on the patenting activity of Norwegian firms, and Czarnitzki et al. (2011), who find a positive effect on the introduction of new products in a sample of Canadian firms.

Finally, the theoretical part of this chapter contributes to a strand of literature that discusses the role and optimal design of innovation policy and to the small but growing literature on inefficiency in the direction of research. Acemoglu et al. (2014) present a model in which directed R&D subsidies are used to transition from dirty to clean technologies. Akcigit et al. (2013) find that subsidies should be directed at basic research that has high spillover effects. Besanko and Wu (2010) compare subsidy policies in an exponential bandit model where a decision maker decides when to give up on a project that has social value. Bryan and Lemus (2016) and Hopenhayn and Squintani (2016) present models in which strategic interaction leads to over-exploitation of incremental or simple projects. In contrast, in this chapter, I compare the solution of the decision problem faced by a single firm with the solution preferred by the social planner, assuming a particular relation of private and social values. Bryan and Lemus (2016) emphasize that optimal R&D policies must condition on the properties of inventions not discovered in equilibrium; I assume that the properties of all research projects are known to the social planner. Admittedly, whether or not firms can reveal the properties of all projects that are feasible to a social planner is an important consideration, but outside the scope of this chapter. More tenuously related is Jovanovic and Rob (1990), who present a model in which a decision maker chooses between exploration and refinement of known directions of research.

1.3 Institutional setting and data

1.3.1 The Austrian Research Promotion Agency FFG

The Austrian Industrial Research Fund FFF ("Forschungsförderungsfonds für die gewerbliche Wirtschaft") was set up in 1967. In 2004, the Austrian Research Promotion Agency FFG ("Forschungsförderungsgesellschaft") was founded that incorporated the FFF. In an evaluation study from 2003, conducted by a panel of international policy experts, the authors calculated that the FFF accounted for 80 percent of all direct government funding for Business R&D in Austria in 2002 (Arnold et al. 2004). The main vehicle of funding is the "Basisprogramm". The "Basisprogramm" is a project-oriented research grant program with rolling admission, open to all firms of all sectors, that has existed in a relatively consistent form since the early 1990s. Firms apply for funding with proposals for research projects with no formal constraints on research area or topic.[4] All subsequent analysis is carried out only for applications to the "Basisprogramm".

Funding applications include a detailed description of the project and a quote of the total cost. The agency applies a standardized evaluation procedure to all applications. Applications are evaluated on technical and commercial merits. The technical evaluation consists of an assessment of the technical quality of the project and an assessment of how the project relates to the firm's technical capabilities. The commercial evaluation consists of an assessment of the commercial value of the project and an assessment of the commercial performance of the firm. Each of the four assessments results in an integer-valued summary score between

[4] In 2002, the "Basisprogramm" accounted for 80 percent of the total value of funds provided by the FFF (Arnold et al. 2004). The remaining 20 percent are specialized technology programs where, typically, a large share of the applicant pool is comprised of universities. In contrast, applications by universities to the "Basisprogramm" are negligible in number.

0 and 50. The four scores are (in their original German names) the *pwert* for the technical quality of the project, the *twert* for the quality of the project relative to the firm's technical capabilities, the *fwert* for the commercial value of the project and the *wwert* for the commercial performance of the firm. The evaluation of the technical quality of the project, which results in the *pwert* score, proceeds as follows: first, the application is assigned to one examiner. The agency has in-house expert examiners for each technological field. The *pwert* score is the sum of four subscores: "novelty", "risk/challenge", "practical value" and "environmental effects". Every subscore is based on 2-5 sub-subscores that each in turn pertain to a different technical characteristic of the project. Every sub-subscore takes a value on a scale from 0 to 4. This scale has an ordinal, but not cardinal, interpretation from worst to best. For each sub-subscore, the examiner has to document the search and review process and provide a reason for the assigned value. Based on the respective sub-subscores, the examiner awards subscores. Every subscore has a discrete grid of admissible integer values that does not contain all values between 0 and the maximal attainable subscore. The rules that map sub-subscores into subscores are commonly understood by all technical examiners. The four subscores are then added up and give the *pwert* score. Because subscores are placed on discrete grids of admissible integer values, some *pwert* scores are more frequent than others, simply because for some scores, there are more possible combinations of subscores that lead to the same *pwert* score than for others. However, in principle, with the exception of the range of 1-9 points, all values between 10 and 50 are attainable. *Pwert* scores in the range of 1-9 points are not attainable for the following reason: in order for a project to score in this range, it must have received at least one subscore equal to 0. Receiving at least one subscore equal to 0 results in the *pwert* score being censored at 0. According to the interviewed examiner, a *pwert* score of 0 eliminates the project as a candidate for funding.[5] There are

[5]However, there is one project in the sample with a *pwert* score of 0 that was approved eventually. In all likelihood, this was one instance were the board of the agency overruled the expert assessment. See discussion further below.

also projects with a *pwert* score of 0 that would otherwise have received scores above 10, had they not received a "knock out" 0 on at least one subscore.

When the first examiner is done, a second examiner carefully reviews the first assessment. If the second examiner disagrees with one sub-subscore, the two examiners discuss and revise the contested item. During my interview with an examiner, it was emphasized to me that sub-subscores are the essential unit of analysis and that aggregates of the sub-subcores, i.e. the subscores and the *pwert* score, are never the subject of discussion during the technical examination. After the two examiners have agreed on all sub-subscores, the sub-subscores are again aggregated into subscores and then further into the *pwert* score. By requesting documentation for each individual sub-subscore and by focussing the attention of examiners on sub-subscores, not on aggregates, the agency is deliberately trying to deter arbitrary and strategic "point nudging" with the *pwert* score.

The assessment of how the project relates to the technical capabilities of the firm is done after the *pwert* score is settled.[6] The commercial assessment follows a similar methodology. Since the *pwert* score plays a distinctively important role in the approval process and is most predictive of the eventual decision, my analysis focuses on the *pwert* score.[7]

[6]This assessment results in the *twert* score, which is the sum of the subscores "increase in know-how", "R&D dynamics" (which is based on sub-subscores that measure the increase in scope of research among other things) and "feasibility", which measures the firm's R&D capability. The *twert* score is assigned jointly by all technical examiners of the respective technological field. The procedure is similar to how the *pwert* score is assigned: the team of examiners works sequentially through all sub-subscores and aggregates. Overall, the *twert* score is considered a less detailed assessment than the *pwert* score.

[7]The funding approval decision exhibits pronounced discontinuities in the *pwert* score. These discontinuities, which are used to identify the causal effect of the funding approval decision, are markedly more pronounced for the *pwert* score when compared to the other scores (discussed in section 1.4 footnote 19). The *pwert* score has other desirable properties: it is determined independently for each application. In particular, the *pwert* score is independent of the identity of the submitting firm (in the sense that, if the same project was submitted by two different firms, they would receive identical scores), whereas *twert* and *wwert* scores are firm-specific. Therefore, the *pwert* is less correlated with firm characteristica than the firm-specific scores and there is less correlation between the scores of different applications by the same firm. The commercial evaluation scores are determined differently depending on the type of firm and are therefore

The decision whether or not a funding application is approved rests with a board composed of six members from the statutory employer association, four members from statutory labor unions and one member from the statutory agricultural association. The board holds between 9 and 10 meetings in a year and handles around 120 applications in each session. In preparation for the board meeting, technical and commercial examiners discuss the current batch of applications. Based on the combination of *pwert*, *twert*, *wwert* and *fwert* scores, the examiners allot projects to the categories "recommended for funding", "not recommended for funding" and "to be discussed". There is no formal rule that stipulates how the scores are to be weighted in arriving at a funding recommendation decision and there are no hard cut-offs for individual scores. However, according to the interviewed examiner, projects with a *pwert* score below 23 points are put into the category "not recommended for funding", unless they received high *twert*, *wwert* and *fwert* scores. Projects with a *pwert* score between 23 and 25 points are very often left "to be discussed" and projects with a higher *pwert* score are usually "recommended for funding", unless they received low *twert*, *wwert* and *fwert* scores. The examiners discuss the applications until unanimous consent is reached. Projects in the category "to be discussed" are borderline cases. The board receives the expert assessments, descriptions of the projects and the funding recommendations, which are not binding for the board. In rare cases, the board overturns the funding recommendation (in either direction). The board makes funding decisions on the projects in the category "to be discussed". A wide range of considerations, including industrial policy and the relationship with the companies, potentially influence board decisions. The exact rules under which the board operates and the transcripts of the board meetings are confidential. The agency does not disclose the identity of successful and unsuccessful applicants to the public.

somewhat ambigious. I discuss robustness checks that utilize the other scores in section 1.5.

Once an application is approved, the funding amount scales with no (absolute) upper limit to the total cost of the project. On average, applicants receive around 25 percent of the total cost of the project in cash ("free money"), and in some cases, additional funds in the form of loans or guarantees. Still, cash payments are the main funding instrument of the agency and accounted for around 85 percent of the overall cash value of funding in 2002 (Arnold et al. 2004).[8] The average funding amount was 153,000 Euro across all applications that were approved between 2002 and 2005. The funding does not come with special contractual provisions regarding patenting. Firms that receive funding are not required to patent the offspring of their research efforts and the creation of patents is not an explicit policy goal. Before I collected the data, the agency did not have patent records of the firms they were funding.

1.3.2 Data and construction of the baseline sample

The data used in this chapter stems from three sources: the funding agency FFG itself, the patent database PATSTAT by the European Patent Office and the database AMADEUS by Bureau Van Djik. The data from the funding agency FFG covers all applications to the "Basisprogramm" between 2002 and 2005. For most applications, data is available on the total cost of the project in Euro, as specified by the firm upon submission, the evaluation

[8]The matching rate of a project, defined as the funding amount divided by the total cost of the project, depends to a limited extent on firm and project characteristics. A relatively small fraction of applications, either for extremely risky projects, or from financially constrained small firms, received considerably higher matching rates. In general, the agency regards the total cost of a project as an exogenous property and aims to provide 25 percent of the cost. In compliance with EU guidelines, the agency has gradually shifted over time towards discriminating more between small and large firms in its funding policy. However, according to the agency, this was less the case between 2002 and 2005. Figure 16 in the Appendix depicts the distribution of the matching rates for funded projects in the baseline sample. One reason why matching rates vary in practice is that the agency applies uniform accounting rules to projected wages and capital rental rates (in contrast to the accounting rules established by the firms). In the Appendix Figure 17, I provide plots of the matching rates for funded projects against the *pwert* score. There is no evidence that matching rates for funded projects depend on the *pwert* score. According to the agency, the total cost of a project does not enter the evaluation process in a direct way.

scores, the year in which the application was received, the funding decision and the funding amount. Furthermore, for many applicant firms, the FFG provides balance sheet data on sales, R&D expenditures, employment, exports and cashflow for the last three years preceding the funding application.[9]

PATSTAT is the most comprehensive data source for patent applications to the European Patent Office and all National Patent Offices in Europe.[10] I manually match the set of Austrian FFG applicant firms to the set of patent applicants with an Austrian country code in PATSTAT using the firm's name and address. The fact that Austrian law encourages splitting all areas of operation of a firm into distinct legal entities facilitates the assignment of patents on a decentralized level.[11] I aggregate patent information into patent statistics per firm and year due to concerns about the effectiveness of anonymization by the agency.

The last data source AMADEUS is a financial database maintained by Bureau Van Djik that contains information on various balance sheet items of private firms, along with information on the industrial sector of operation and firm age. I manually match the set of Austrian FFG applicant firms to the set of firms in AMADEUS that have an Austrian country code using the firm's name and address. While financial data is missing for most firms in the sample, the AMADEUS matching allows me to assign FFG applicant firms to sectors based on the NACE2 classification. In addition, it provides the year of incorporation for the

[9]The FFG does not follow up on balance sheet data after the funding application. This means, for example, that it is not possible to study the effect on R&D expenditures with data from the agency alone.

[10]PATSTAT contains information from the patent record, including application date, forward and backward citations, and technology class. Importantly, PATSTAT includes information on patent families, which are the sets of patents protecting the same invention across several jurisdictions.

[11]Firms are split into distinct legal entities to limit liabilities across branches. Unlike most other countries, Austria explicitly permits shifting profits across affiliates in a tax efficient way. Hence, there is no downside to decentralizing. It is probably more appropriate to consider this an assignment on the establishment level. Whenever firms are owned by no more than one private individual, I also match the patents of the owner to the firm.

majority of the firms.[12]

The construction of the sample used for the analysis involves some exclusions from the full set of grant applications. In my research design, I relate patent applications by the firm after the funding application to the binary funding approval decision, using the *pwert* score as an instrument. I exclude applications from firms with multiple funding applications in the same year. This is an important caveat and shortcoming of this chapter, since I exclude some of the biggest firms and the most frequent recipients of FFG research grants. The remaining applications constitute 62.6 percent of the total number of applications, and the remaining applicant firms constitute 92 percent of the total number of applicant firms. Table 29 in the Appendix shows the number of applications per firms in each year between 2002 and 2005. After eliminating applications that are missing the *pwert* score or the cost of the project, the final sample consists of 2619 funding applications across 1936 firms.[13] The applications are pooled across years. For inference, I cluster the standard errors of applications belonging to the same firm.

For more than one third of the applications in the baseline sample, firm-level data by the agency on R&D expenditures, employment, sales and cashflow over the pre-application period is missing. This data is missing in a systematic way: applications that were rejected and applications with low evaluation scores are more likely to have missing data. Therefore, I drop these variables from the baseline sample for my analysis. I present descriptive statistics for the available data on employment, R&D expenditures, sales, firm age and patent applications for firms in the baseline sample in Table 2. However, these numbers may be of

[12]See the Appendix for more information on the matching and how missing values were imputed.

[13]For around 10 percent of the remaining funding applications, evenly distributed across years, the *pwert* score is missing. Applications with missing *pwert* score are mainly "desk-rejects", which are very obviously without chance for funding. For two more applications, the total cost of the project is missing.

limited international comparability. As mentioned above, Austrian companies are decentralized for legal and fiscal reasons and most firms report balance sheet data for each affiliate individually. The US SBIR program, in contrast, has a firm size constraint that encompasses all affiliates.

1.3.3 Variables in the baseline sample

The baseline sample has 2619 observations, each corresponding to exactly one distinct funding application, across 1936 firms. For each funding application, the *pwert* score, the funding approval decision, the total cost of the submitted project, patent applications by the firm before and after the funding application, sector of the firm, firm age and year of application are available.

The *pwert* score is, as discussed in the preceding sections, the evaluation score for the technical quality of the project, taking integer values between 0 and 50. The full distribution of *pwert* scores in the baseline sample can be found in the Appendix Table 30. *Approved* is an indicator variable that takes on the value 1 if the funding application was approved. 66 percent of all funding applications in the baseline sample were approved. *Project cost* is the total cost of the submitted project as quoted by the firm in hundred thousands of Euro. Table 1 presents descriptive statistics of the cost of submitted projects and funding amounts. The average funding amount (conditional on funding approval) is 118,000 Euro in the baseline sample, the single largest amount being 2,300,000 Euro. The outcome variables *EP Patent binary post* and *Nat Patent binary post* are indicator variables that measure whether or not the firm filed an application for a (for at least one) European Patent or a National Patent (with a patent office other than the EPO) in the year of the funding application or during the subsequent three years after the funding application. To be clear, for patent applications filed in the same year as the funding application, it is not known whether the patent application was filed before or after the firm applied for funding.[14] The variables *EP Patent binary pre* and *Nat Patent binary pre* measure whether or not the firm had filed patent applications in the period preceding the funding application. The variables *EP Patents pre* and *Nat Patents pre* are the corresponding (integer-valued) count variables. These variables, capturing the firms' patenting histories, are truncated at 12 years before the funding application. Descriptive statistics of patent variables are presented in Table 2. A description of the dummy variables for the sector of the firm and firm age can be found in the Appendix.

The distinction between European Patents and National Patents pays testament to the institutional environment. When filing a patent application, an Austrian firm has to choose the jurisdictions where it wishes to protect its invention. This choice reflects in part the perceived value of the invention.[15] The Austrian market with a total population of around 8 million is small. With an application to the European Patent Office, the firm can protect its invention in all of the EU and the rest of the EPO member states.[16] Alternatively, the firm can directly file patent applications with national patent offices. The set of patent applications that pertain to the same invention is called a "patent family".[17] I partition patent families into those that comprise an application to the European Patent Office, which I call "European Patents", and those that do not, which I call "National Patents". In practice, most Austrian firms either file a patent application with the Austrian Patent Office only (which is a National Patent according to my definition) or they file a patent application with the European Patent Office.[18]

[14]In section 1.5, I show that the treatment effect on patent applications in the year of the funding application is negligible and that the effect has considerable delay.

[15]For example, obtaining a European Patent is considerably more expensive.

[16]The European Patent must be validated in the states where the applicant wishes to protect the invention.

[17]I use the PATSTAT DOCDB definition of a patent family in this chapter.

[18]For 50 percent of the European Patents in the sample, the firm first filed the application with the Austrian Patent Office, and later filed the application with the European Patent Office. When a firm files

1.4 Empirical strategy

The dependence of the probability of funding approval on the evaluation score *pwert* is plotted for the baseline sample in Figure 1. The plot reveals a very steep, almost discontinuous, increase in the point range of 23 to 28. Since the evaluation score *pwert* is a sum of separate subscores (described in section 1.3), it is reasonable to assume that the dependence of project quality on *pwert* score does not exhibit the same steep increase. Therefore, the distinct steep increase in the funding approval probability can be used as an instrument for the effect of funding approval.[19]

This empirical strategy is more likely to produce credible estimates of the causal effect of funding approval than the OLS estimate obtained from regressing patent outcomes on funding approval. The OLS estimate partly reflects the selection policy of the government agency and may overestimate or underestimate the causal effect. If the agency systematically "rewards winners", the OLS estimate may overstate the effect of the research grant. On the other hand, if the agency is trying to "lend a helping hand" to firms that struggle or face particularly great challenges in research, the OLS estimate may have downward bias. The downside to my empirical strategy is that the estimated effect has a local interpretation,

---

an application for a specific invention with the Austrian Patent Office, it obtains a "priority date". The firm reserves the right to file a patent application for the same invention with the European Patent Office within one year from the priority date. In this case, I regard the date of the patent application to the Austrian Patent Office as the filing date of the European Patent.

[19] In section 1.3, I discuss how the *pwert* score is determined. The score has several desirable properties (in contrast to the other evaluation scores, discussed in section 1.3 footnote 7) and plays a distinctively important role in the funding approval decision. The quality of the non-linear increase as instrument depends on how steep the increase is. The steeper the increase, the higher the fit in the first stage regression. The other evaluation scores, the firm-specific technical evaluation score *twert*, the firm-specific commercial evaluation score *wwert* and the project-specific commercial evaluation score *fwert*, achieve considerably lower R^2 than the *pwert* in the first stage, in declining order, and exhibit less pronounced jumps. I use the *pwert* throughout the chapter for my analysis and perform robustness checks with the *twert* as instrument. The probability of funding approval is plotted against *pwert* and *twert* scores in the Appendix Figure 18. Furthermore, I include the other evaluation scores as controls as a robustness check. The downside of using the other evaluation scores as controls is multicollinearity between scores. Since the *pwert* score is used to construct the instrument, multicollinearity decreases the precision of the estimate.

i.e. applies to projects in the point range of the steep increase in funding approval probability.

The estimated econometric model is a two stage model similar to a Fuzzy Regression Discontinuity Model. In the first stage, the dependence of funding approval probability on $pwert$ is modelled with linear splines on the ranges $[0, p^*]$ and $[p^*, 50]$ and a steep, non-linear increase starting at p^*. I model the steep increase starting at p^* in different ways and select the location of p^* from data by best fit in the first stage. In the first model, the steep increase is modelled as a cubic spline starting at $pwert = 21$. Two alternative models, where the increase is modelled using two or three discrete jumps respectively, are considered as main robustness checks.[20] The linear spline function starting at p^* is defined as

$$l_{\geq p^*}(pwert) = max(pwert - p^*, 0)$$

In the second stage, I control for the direct dependence of the outcome variable on $pwert$, i.e. the dependence of project quality on score, with linear splines on $[0, p^*]$ and $[p^*, 50]$ only, without the steep, non-linear increase at p^*. For identification, it is necessary to assume a less flexible functional form for the direct dependence of the outcome variable on the evaluation score $pwert$. The two regression equations of the model that uses the cubic spline as instrument are given by

$$\begin{aligned} approved_i = \alpha + \beta_1 pwert_i + \beta_2 l_{\geq 21}(pwert_i) \\ + \beta_3 l^2_{\geq 21}(pwert_i) + \beta_4 l^3_{\geq 21}(pwert_i) + \gamma X_i + \epsilon_i \end{aligned} \tag{1}$$

$$EP\ Patent\ binary\ post_i = \alpha' + \tau\ \widehat{approved}_i + \beta_1' pwert_i + \beta_2' l_{\geq 21}(pwert_i) + \gamma' X_i + u_i \tag{2}$$

[20] In the second model, the increase is modelled using discrete jumps at $pwert = 23$ and $pwert = 27$. The third model has discrete jumps at $pwert = 23$, $pwert = 26$ and $pwert = 29$.

with orthogonality condition $l^2_{\geq 21}(pwert_i), l^3_{\geq 21}(pwert_i) \perp u_i$, where X_i is a vector of control variables. The model uses variation in a narrow point range of the *pwert* score that produces big differences in funding approval probabilities to identify the effect. It does not use variation in funding approval decisions. Projects with a *pwert* score of 20 points that were approved are innately better projects than rejected projects with the same *pwert* score (see section 1.3 for a description of the award process). The board may have had good reasons to prefer one project or one firm over another, even though they may have received identical scores by the examiners.

Figure 1 shows the fit of the functions estimated in the first stage for all three models. The estimation strategy is transparent. For the model that uses the cubic spline as instrument, and for the model that uses three discrete jumps as instruments, the left column in Figure 2 shows how the binary outcome variable *EP Patent binary post*, which takes on value 1 if the firm filed a (at least one) patent application for a European Patent after the funding application, is fitted in the second stage regression (without covariates). The curve in red depicts the predicted value of *EP Patent binary post* omitting the estimated treatment effect of funding approval, while the curve in black depicts the predicted value of *EP Patent binary post* with the estimated treatment effect. To be more precise, the red curve shows the estimated direct dependence of *EP Patent binary post* on the evaluation score *pwert* in the second stage given by

$$\hat{\alpha}' + \hat{\beta}_1' pwert + \hat{\beta}_2' l_{\geq 21}(pwert)$$

The black curve is the sum of the red curve and the predicted propensity to be approved multiplied by the estimated treatment effect of funding approval given by

$$\hat{\tau}\ \widehat{approved}(pwert) + \hat{\alpha}' + \hat{\beta}_1' pwert + \hat{\beta}_2' l_{\geq 21}(pwert)$$

The non-linear increase in the funding approval probability "lifts" the linear spline. In the middle column of Figure 2, I present the second stage fit when I re-estimate both models with observations $pwert = 0$ excluded (see section 1.3 for a discussion of how to interpret $pwert = 0$). In the right column, I present the fit for a different model specification, in which I exclude the linear spline starting at p^* from the second stage and control for the *pwert* score linearly (observations with $pwert = 0$ are again excluded).

Specification of other control variables

The set of control variables comprises past European Patent and National Patent filings, project cost and fixed effects for firm age, firm sector and year of application. Past National Patent filings and European Patent filings are included seperately as indicator functions and quadratic functions of the patent count. These variables capture "pre-sample" patent information and control for the firms' fixed propensities to file patents, in the vein of Blundell et al. (1999). Project cost are included as a quadratic function. For further explanation of the variables see section 1.3.2.

1.5 Results

In section 1.5.1, I present the main results on the effect of government research grants on the propensity to file a European Patent. Section 1.5.2 presents additional evidence that motivates the model studied in section 1.6.

1.5.1 Main results on the effect of government research grants on the propensity to file a European Patent

Table 3 presents the results for the effect of funding approval on the propensity to file a (at least one) European Patent in the year of the funding application or during the subsequent three years. I discuss the estimates from the model that uses the cubic spline as instrument (columns 1 and 2), noting that the results are very similar for the models that use two or three discrete jumps as instruments (columns 3 to 6). The point estimate of the effect of funding approval is 12.8 percentage points in the parsimonious model without control variables (including only controls for *pwert* score), which is statistically significant at the 0.05 level (column 1). When I include the full set of control variables, the point estimate of the effect of funding approval is revised by 0.35 standard deviations to 10.8 percentage points (column 2, statistically significant at the 0.05 level). I regard this estimate as my main result. The lowest point estimate across all models is 10.1 percentage points. The mean of the dependent variable is 27.8 percentage points in the sample. Figure 2 illustrates the result graphically (explained in section 1.4). The results for the first stage regressions are shown in the Appendix Table 31.

Validity checks

Table 5 contains the results from the placebo test of applying the IV model to the 4 years preceding the funding application. In the upper panel of Figure 21 in the Appendix, I plot the average propensity to file a European Patent against the *pwert* score for both periods, the 4 years preceding the funding application and the 4 years following (including the year of) the funding application. While the plots look similar, there is no statistical evidence that European Patent filings in the 4 years preceding the funding application exhibit a non-linear increase in the *pwert* score correspondent to the non-linear increase in funding approval probability. Plots for all control variables against the *pwert* score and regression results from applying the IV model to all control variables can be found in the Appendix.

Robustness checks

The results appear robust to a variety of alternative specifications, like excluding applications that received a *pwert* score equal 0 (see section 1.3.1 for a discussion of why such applications are ambiguous), controlling for *pwert* score linearly (instead of using linear splines) and restricting the sample to observations immediately before and after the steep increase in the funding approval probability. Using the *twert* score to construct my instrument and using the other evaluation scores as controls yields very similar estimates.[21] To account for the uncertainty that the break points of linear splines and the locations of jumps are chosen by best fit from data, I implement a bootstrap in which I first re-sample firms with replacement and then choose break points and jumps by best fit for the obtained sample. All robustness checks can be found in the Appendix.

Time delay of the effect and results for extended time frame

The results are presented with patent outcomes measured for the year of the funding application and the subsequent three years. When I extend the time frame over which patent outcomes are measured by one year, the results become stronger. The estimate of the effect of funding approval on the propensity to file a European Patent in the year of the funding application or during the subsequent 4 years is around 12 percentage points (Appendix Table 33). I examine the delay of the treatment effect by studying the effect of funding approval on consecutive two-year periods after the funding application. Figure 3 shows point estimates and confidence intervals. I find the highest point estimate of 9.9 percentage points,

[21] Due to multicollinearity between scores, these estimates are less precise than the estimates in the main specification.

statistically significant at the 0.05 level, for years 3 and 4 (year 0 being the year of the funding application). In contrast, the effect of funding approval is only around 4 percentage points and statistically insignificant for the years 1 and 2. Years 0 and 1 yield even weaker results. After year 4, the estimated treatment effect decays and is statistically indistinguishable from 0 ever after.

A potential concern could be that the effect is explained by firms feeling "obliged" to file patents (on inventions that otherwise simply would not be patented) in order to obtain tangible proof of success for the agency. Such an explanation is at odds with the finding that the effect has considerable delay. The grant is paid out within one year after funding approval (see section 1.3). The agency does not follow up with the firm after the funding contract is fulfilled. Furthermore, as mentioned in section 1.3, the agency does not regard the creation of patents an explicit policy goal and patents are not a requisite for future applications.

Discussion of OLS estimates

The estimates obtained using my empirical strategy contrast with the estimates obtained from a OLS regression of the propensity to file a European Patent on funding approval, presented in Table 4. The OLS estimate of the effect of funding approval in the model without controls is 17.4 percentage points, with a standard deviation of 1.82 percentage points. When I include the full set of control variables in the OLS model, the coefficient of funding approval drops by around 6.5 standard deviations. Columns (3) and (4) in Table 4 show the placebo effect of funding approval on European Patent filings in the 4 years preceding the funding application in the OLS model. Even controlling for lagged patenting histories in the placebo regression, funding approval is highly correlated with European Patent filings in the 4 years preceding the funding application, statistically significant at the 0.01 level. This suggests that the OLS estimates, in all likelihood, do not capture the causal effect. Persistent heterogeneity in firm characteristics, already present in the 4 years preceding the funding application, is a likely source of bias.

Heterogenous treatment effects by firm age

Figure 19 and Figure 20 in the Appendix show the distribution of firm age in the baseline sample. The age of the applicant firm at the time of the funding application is available for 2444 applications (93.3 percent of applications) in the baseline sample. The most plausible reason for why the age may be missing is that the applicant never incorporated. For 10 percent of the applications in the sample, the firm age is negative, which means that the firm incorporated after the funding application. The FFG requires applicants that receive funding to incorporate.[22] There is some noise in the measurement of the firm age. If a firm changes its legal form, it re-incorporates and thereby resets the official year of incorporation. To the best of my knowledge, it is not possible to distinguish between genuinely new firms and firms that are formed as the result of a restructuring.

I count firms that are missing the year of incorporation as firms of age 0 and split the sample at the median age, which is 5 years. I refer to firms above the median age as "established firms", and to firms below the median age as "young firms". Since firms that are missing the year of incorporation in all likelihood never incorporated, it is reasonable to regard such firms as "not established as of the time of the funding application". As a robustness check, I also consider the split sample estimates after excluding firms that are missing the year of incorporation. The results are presented in Table 6. For the model that uses the cubic spline as instrument, the point estimate of the effect of funding approval is 17

[22]In principle, legal incorporation is necessary when taxable income is earned or when employees are hired. It is also worth pointing out that some firms file a patent application before they incorporate: in this case, the inventor (or the inventor team) applies for the patent and has the patent assigned to the newly founded firm later.

percentage points for established firms, statistically significant at the 0.01 level, compared to 3.63 percentage points for young firms, which is not statistically significant. The mean of the dependent variable is 33.2 percentage points for established firms and 22.9 percentage points for young firms. Across all models, the treatment effect for young firms lies between 1.85 and 2.2 standard deviations below the treatment effect for established firms. The result is robust to excluding applications that received a *pwert* equal 0 and controlling for the score linearly, instead of using linear splines. Excluding firms that are missing the year of incorporation and splitting the sample at the new median age, which is 6 years, produces very similar results. All results can be found in the Appendix. Results for the placebo test of applying the model to the 4 years preceding the funding application can be found in the Appendix Table 36. Hence, the findings from this section suggest that grants to established firms have larger effects.

Results on the propensity to file a National Patent

Table 32 contains the results for the effect of funding approval on the propensity to file a National Patent in the year of the funding application or during the subsequent three years. In the model that uses the cubic spline as instrument and a full set of controls, the point estimate of the effect of funding approval is -0.1 percentage points. The models that use two and three discrete jumps as instruments yield point estimates of -0.7 percentage points and 1.5 percentage points respectively. The effect of funding approval is not statistically significant in any model. The mean of the dependent variable is 25 percentage points.

1.5.2 Additional evidence that established firms undertake ambitious projects due to government research grants

Survey evidence

It is evident that the institutional features of the FFG, in particular the funding and selection rules described in Section 1.3, are important in explaining the effect of the grants. The Austrian Institute of Economic Research WIFO, together with numerous international innovation policy experts, conducted an evaluation study of the FFG in 2003. The reports delivered detailed insight into the structure and processes of the organization. In a review of the funding instruments, the authors describe the mission of the FFG and express concerns that the agency may not be effective in meeting a central policy goal:

> "The way FFF tailors its funding instrument complies with the normal funding rationale: FFF rewards high risk. What remains questionable however is the way FFF deals with the second innovation barrier: access to financial resources." (Arnold et al. 2004, p.37)

The authors acknowledge that different types of innovation barriers are relevant for different types of firms. While they seem generally content with how the agency promotes "more risky" research among established firm, they suggest raising the cash value of subsidies for financially constrained young firms. My findings offer some evidence that the FFG may indeed be more effective in addressing innovation barriers faced by established firms. As part of the evaluation, the WIFO conducted a large scale survey among rejected and successful applicant firms. The authors summarize the survey responses by rejected firms as follows:

> "Our survey suggests that just under a third of firms cancel projects if FFF rejects their application for funding. [..] Of the two-thirds of firms that carry on with projects without FFF funding

- 32-43% delay the start of the project
- 51-61% do the project over a longer period of time
- 60-74% reduce the size of the project
- 40-49% make the project technically less ambitious
- 63% get the results of the project later than originally planned

(Arnold et al. 2004, p.53)

The range of reported responses reflects differences between firms of different size. Furthermore, the authors stress that

> "Interestingly, it is not the largest companies [..] who are least affected. Very large firms generally have a large supply of alternative projects at any time, so FFF may be more decisive for which project gets done, rather than for whether any R&D is done at all." (Arnold et al. 2004, p. 52)

In the survey with successful applicant firms, the firms were asked to compare FFG-funded projects with internally financed research projects. Around two-thirds of the successful applicant firms responded that FFG-funded projects were technologically more challenging and difficult than internally financed projects. By number of employees, the fraction of firms who found FFG-funded projects more difficult was 51.1% (0-9 employees), 63.8% (10-99 employees), 63.4% (100-249 employees), 71% (250+ employees). Putting the evidence from successful and rejected applicants together, the authors conclude that

> "In many cases, FFF funded projects represent a step towards riskier projects for the individual firm, irrespective of whether the projects are risky in a more objective [in an absolute] sense." (Arnold et al. 2004, p. 55)

And further,

> "It is clear that a FFF subsidy brings significant value to the Fund's beneficiaries. It prompts companies to perform additional R&D. In a third of the cases, it allows projects to be performed that would otherwise not have been undertaken, and in most of the remaining cases it allows projects to be bigger, quicker and to take more technical risks that would otherwise be the case, allowing companies to bring improved products and processes faster to the market." (Arnold et al. 2004, p. 57)

Evidence from patent measures

In this section, I present evidence that is consistent with the notion that research grants induce established firms to undertake ambitious research projects, using a variety of patent measures. First, I show that marginal patents produced by established firms due to grants appear non-trivial with respect to commercial value, conventionality and quality. Second, I show that for firms with prior patent filings, funding approval is correlated with heavier utilization of knowledge novel to the firm.

For the first set of results, I focus on firms above the median age in the sample, which is 5 years. In section 1.5.1, I found that for this group of firms, a research grant increases the propensity to file a European Patent within 4 years by around 17 percentage point. On the other hand, there is no discernible effect on the propensity to file a National Patent.[23] Since European Patents have greater geographical coverage, they are more commercially valuable. I then partition the patents in my sample based on two different indicators. The first indicator is a patent conventionality measure which was recently suggested by Berkes &

[23] In section 1.5.1, I discuss that the estimated effect is statistically insignificant and negative in the entire sample. The same holds true for both, firms above and below the median age when I split the sample.

Gaetani (2016).[24] The second indicator is a standard measure of patent quality, the number of citations received by a patent.[25] I rank all patents in PATSTAT by applicants from 10 different European countries by their conventionality and by the number of citations received.[26] Patents that are more unconventional than the median patent filed in the same year and technological class are classified as "unconventional".[27] Similarly, patents with a higher citation count than the median patent filed in the same year are classified as "high-quality" patents.[28] The point estimate of the effect of a research grant on the propensity to file an unconventional patent for firms older than 5 years is around 16 percentage points, while the estimate of the effect on the propensity to file a conventional patent is only around 2 percentage points (Table 7). Similarly, I find that the estimated effect is around 13 percentage points for high-quality patents, but only around 4 percentage points for low-quality patents (Table 7). Hence, marginal patents produced due to grants by firms older than 5 years appear non-trivial with respect to commercial value, conventionality and quality. It is noteworthy that this only applies to the patents produced *due to grants*: it is not at all uncommon for firms in my sample to file conventional patents, low-quality patents or National Patents. The overall propensities to file conventional patents, low-quality patents or National Patents in the 4 years following the funding application are only slightly lower than the propensities to file unconventional patents, high-quality patents or European Patents.

---

[24]It is based on the "unconventionality" of combinations of technological classes cited by a patent. If, for example, a patent cites one patent in IPC class C12 biochemistry and another patent in IPC class A43 footwear, it is considered unconventional because the combination of biochemistry and footwear as knowledge inputs is uncommon. More details of the procedure are described in Appendix D.

[25]If a patent is cited by many other subsequent patents, it is regarded to be of great quality. See, for example, Trajtenberg 1990 and Hall et al. 2005.

[26]In many instances, there are not enough Austrian patents by year or technological class to base the ranking just on Austrian patents.

[27]"Technological classes" correspond to IPC-classes.

[28]I refrain from normalizing the citation count by IPC class to capture the fact that firms may move into "hotter" technological areas. Results when I normalize by IPC class are generally weaker.

For the second set of results, I compute an aggregated firm patent measure that captures one aspect of technological risk and ambition, namely the average utilization of technological knowledge novel to the firm. This measure, which I call "Average Firm Novelty", compares the technological classes cited by new patents with the technological classes cited by all previous patents filed by the same firm.[29] For each firm-application pair, I record all distinct technological classes that were cited by any patent filed by the firm in the year of the funding application or during the subsequent three years. Then I count how many of these technological classes had not been cited by any patent filed by the same firm in the pre-application period (between 1980 and the year before the funding application). This number is then normalized by the total number of patents filed in the year of the funding application or during the subsequent three years.[30] For example, for firms applying for funding in 2002, it is defined as follows:

$$avg\_firm\_novelty = \frac{\#\{\text{Novel technological classes cited 2002-2005}\}}{\#\{\text{Patents filed by the firm 2002-2005}\}}$$

where

$$\{\text{Novel Technological classes cited 2002-2005}\} =$$

$$\{\text{Technological classes cited by any patent of the firm filed between 2002 and 2005}\}$$

$$\setminus \{\text{Technological classes cited by any patent of the firm filed between 1980 and 2001}\}$$

The measure has a caveat: it is only meaningful for firm-application pairs with observed patent filings before and after the funding application. The subsample of applications that satisfy these criteria contains 686 funding applications across 485 firms, of which 537 were

[29] A "technological class" corresponds to a IPC class.

[30] The normalization by the number of patents filed by the firm after the funding application is essential since otherwise, an increase in "Average Firm Novelty" could spuriously be caused by an increase in the total number of patents, with some of the additional patents randomly citing technological classes novel to the firm.

approved and 149 were rejected.[31] Due to the small sample size, I refrain from using the two-stage model of section 1.4 and instead compare approved and rejected firms. The results of this descriptive regression can be found in Table 8. Firms that received funding cited on average 0.2 more technological classes per patent that were novel to the firm than rejected firms in the year of the funding application or during the three subsequent years. However, with a standard error of 0.151, this difference is not statistically significant. Controlling for firm age, firm sector and year of application fixed effects, the estimate of the difference is revised to 0.302 technological classes per patent that were novel to the firm, which is statistically significant at the 0.05 level. The revision is due to the inclusion of sector fixed effects, which suggests that more funding was channeled towards conservative sectors. In the sample, firms cited on average 3.41 distinct technological classes per patent in the year of the funding application or during the subsequent three years, of which 1.16 technological classes were novel to the firm.

1.6 The Model

In this section, I interpret the findings from section 1.5. The finding from section 1.5.1, that the effect of the grant is weaker for younger firms, is inconsistent with an explanation whereby FFG research grants mainly affect firms through alleviating financial frictions. The evidence from section 1.5.2 is inconsistent with an explanation whereby marginal projects undertaken by established firms due to the research grants are of poor quality and, for this reason, would not be undertaken in the absence of a grant.

In my model, I assume that a risk-neutral firm starts off as "young" and becomes "established" upon successfully exploring a research line. After an initial "breakthrough", the

[31] For 689 firm-application pairs patent filings before and after the funding application are observed. For three more firm-application pairs, the technological classes cited by the patents filed during the pre-application period are missing.

new research line spawns a series of incremental, safe projects of declining value that can be researched subsequently. Over the next periods, the established firm receives ideas for other research lines that would render the old research line obsolete. The firm faces the choice between exploring a novel research line, thereby potentially replacing its current research line, and exploiting the old research line. If the old research line is below a threshold age, i.e. has not been exploited sufficiently, the firm dismisses the idea for the novel research line. Assuming that the social value of a research project scales linearly to the private value of a research project, I show that the social planner renews research lines more aggressively and starts exploring earlier than the firm. To correct for this inefficiency, the social planner offers a research grant for the risky, explorative project.

The model is consistent with survey evidence from section 1.5.2, that suggests that established firms undertake risky, ambitious research projects when they receive a grant. The model reconciles survey responses that indicate that firms delay or take longer to carry out ambitious projects when their funding applications are rejected. The model clarifies why the proposed mechanism is less relevant for young firms, which explains the findings from section 1.5.1. Finally, the model rationalizes the evaluation rule by the Austrian Research Promotion Agency that favors high-risk projects.

In addition, I use the model to study the relative effectiveness of research grants and undirected tax credits in this setting. I demonstrate that a uniform reduction in the cost of research for both incremental and explorative projects, does encourage more explorative research. The social planner can, in principle, implement the socially optimal research policy through an undirected tax credit. However, the cost of such a tax credit is higher than the cost of an equivalent research grant.

There is one firm in this dynamic decision problem. Initially, the firm is "young" and has no current research line to exploit. In each period, the firm receives an idea for a research line with probability $\tilde{p}$. If the young firm does not receive an idea, it earns a stage-payoff of 0 and time moves forward. If the young firm receives an idea for a research line, it decides whether or not to explore the idea. I assume that $\tilde{p} = 1$ and claim that this has no bearing on the qualitative nature of my results. Exploration costs $c_e > 0$ and yields a success with probability $0 < p_e < 1$. If the young firm decides to explore and succeeds in establishing a research line, it earns a stage-payoff of K and becomes an established firm with a research line of age 1 in the next period. All payoffs are discounted at the rate $0 < \beta < 1$ and there is no terminal period. If I denote the value of a young firm by V_{Young} and the value of an established firm with a research line of age T as V^T_{Estbl}, the Bellman equation of the young firm is

$$V_{Young} = max\Big\{\beta V_{Young}; -c_e + p_e(\beta V^1_{Estbl} + K) + (1 - p_e)\beta V_{Young}\Big\} \qquad (3)$$

A newly established research line spawns a finite number N of incremental projects that can be researched in subsequent periods. I call the number of projects that have already been worked on the same research line the "age of the research line". The next incremental project of a research line of age T delivers payoff $K\lambda^T$, where $\lambda < 1$, succeeds with very high probability $p_i > p_e$ and costs c_i. To facilitate the exposition, I assume that $p_i = 1$ and claim that this has no bearing on the qualitative nature of my results. I assume that all incremental projects have a positive NPV so that even for the last project

$$K\lambda^N > c_i$$

During the subsequent periods, the firm could, in principle, work through the entire series of incremental projects on the current research line. However, the firm continues to receive ideas for yet other research lines (with probability $\tilde{p} = 1$) that offer a novel approach, superior to the old one. As before, exploring a novel research line costs c_e, succeeds with probability p_e and, if successful, fully replaces the old research line. The old research line is assumed to be a fall back. If the firm decides to undertake the explorative project but fails, it can revert to the incremental research project on the old research line. Hence, an established firm with a research line of age T compares the reservation value of exploring the novel research line with the continuation value of the current research line of age T. Consider, for example, a car producer trying to reduce the gas mileage for the next product cycle of a car model: the current technology of the engine can be improved incrementally to reduce gas mileage from 3 g/hm to 2.8 g/hm for the next generation of the model. The car producer is aware of a novel technology that could reduce gas mileage to 2.5 g/hm, but researching this technology is risky. If the car producer decides to explore the novel technology and succeeds, there is no point in undertaking the incremental project on the old technology and all cars in the next product cycle will be equipped with the novel technology. If the car producer decides to explore the novel technology and fails, she can still implement the incremental improvement on the current technology. Alternatively, she could ignore the opportunity to explore the novel technology and immediately undertake the incremental project on the current technology. The Bellman equation is

$$V^T_{Estbl} = max\Big\{K\lambda^T - c_i + \beta V^{T+1}_{Estbl}; -c_e + p_e(K + \beta V^1_{Estbl}) + (1 - p_e)(K\lambda^T - c_i + \beta V^{T+1}_{Estbl})\Big\} \quad (4)$$

If the firm explores and succeeds in establishing the novel research line, it earns stage-payoff K, the old research line is retired and the age of the research line is reset. Equation (4) gives rise to the reservation value condition that exploration is optimal for an established firm with a research line of age T if and only if

$$\frac{c_e}{p_e} - c_i \leq \beta(V^1_{Estbl} - V^{T+1}_{Estbl}) + K - K\lambda^T \tag{5}$$

Exploration is optimal if and only if the gains from renewing the research line outweigh the cost, taking into account the risk of failure. Since both, the continuation value of the established firm and the stage-payoff are decreasing in the age of the current research line, it is clear that there has to be a threshold age below which exploitation is optimal and above which exploration is optimal. It is important to keep the distinction between the age of the firm and the age of the research line in mind. If an old, established firms has a research line that was introduced very recently, it will exploit the research line at first and then try to renew the research line. Young firms on the other hands have no research line to exploit. The reservation value condition that ensures that exploration is optimal for young firms is

$$\frac{c_e}{p_e} \leq \beta(V^1_{Estbl} - V_{Young}) + K \tag{6}$$

Inequality (6) implies that a young firm always has a higher incentive than any established firm to undertake the explorative research project since it has no other research line to fall back on.

For completeness, I also need a description of the case when an established firm reaches the terminal incremental research project N. If the firm does not renew its research line, it becomes an "obsolete" established firm without active research line. I denote the value of such a firm by V^{obs}_{Estbl} and claim that $V^{obs}_{Estbl} = Y_{Young}$. To ensure that some research happens in the firm decision problem, I assume that the total discounted value of a research line is sufficient to motivate some research:

Assumption 1. *The total discounted value of a research line* $K+\sum_{t=1}^{N}\beta^t(K\lambda^t-c_i)$ *is such that*

$$p_e(K+\sum_{t=1}^{N}\beta^t(K\lambda^t-c_i))>c_e$$

I summarize the discussion in the proposition below.

Proposition 1. *Consider the firm decision problem defined above and suppose that Assumption 1 holds. Then,*

- V_{Estbl}^{T} *is strictly decreasing in the age of the current research line* T*. Furthermore,* $V_{Young}=V_{Estbl}^{Obs}<V_{Estbl}^{T}$ *for all* T*.*
- *there exists a threshold age* $\bar{T}$ *such that an established firm with a current research line of age* T *explores the novel research line if and only if* $T\geq\bar{T}$*.*
- *a young firm always undertakes the explorative research project.*

Proof. See Appendix. □

Innovation Policy

I assume that the social value of a research project is a constant multiple of the private value. This means that there is a $S>1$ so that the social payoffs of the projects are SK, $SK\lambda$, $SK\lambda^2$ and so on. In reality, there are multiple possible reasons for why the additional social value would scale to the private value. First, there could be "mismatch" in time-horizons: while private payoffs of an innovation are truncated when imitation occurs, social payoffs are not truncated. If payoffs are discounted at rate β and truncated at time T, S would be given by $1/(1-\beta^{T+1})$. Alternatively, we could imagine that inventions of greater private value also enable more follow-up research by others or deliver more consumer surplus. I define the social planner decision problem as the firm decision problem with all project payoffs scaled by factor S, while costs and discount factor are left unchanged. Obviously, all findings from Proposition 1 are still valid in the social planner decision problem. There is an important equivalent way of looking at the social planner's problem: the social planner's problem corresponds to the firm decision problem with project payoffs left unchanged, but costs lowered uniformly for explorative and incremental projects by the factor $1/S$. Before discussing how the solutions of the two problems relate, I introduce the assumption that the cost of an incremental project does not vastly exceed the cost of the explorative project.

Assumption 2. *The cost of the explorative research project c_e and the cost of the incremental research project c_i satisfy*

$$\frac{c_e}{p_e} > c_i$$

Intuitively, I expect Assumption 2 to be met since $p_e < 1$ and because it seems reasonable to think of the explorative project as being more ambitious and bigger in size. Even if Assumption 2 fails, my results are not invalidated, only the problem is trivialized. As can be seen from the reservation value condition (5), a violation of Assumption 2 implies that exploration dominates exploitation for both, the social planner and the firm. Therefore, to have a non-trivial problem, suppose that Assumption 2 holds. The following proposition shows that the social planner is always more inclined to explore.

Proposition 2. *Consider the firm decision problem defined in the previous section, the social planner problem defined above and suppose that Assumptions 1 and 2 hold.*

Let $\bar{T}$ be the threshold age of the current research line for exploration in the firm decision problem, established in Proposition 1, and let $\bar{T}^*$ be the threshold age of the current research line for exploration in the social planner problem. Then, it must be true that the social planner explores earlier, i.e.

$$\bar{T}^* \leq \bar{T}$$

Proof. See Appendix. □

Hence, the social planner starts exploring earlier and renews the research line more aggressively. Consider an established firm with a research line of age T s.t. $\bar{T}^* \leq T < \bar{T}$. In this case, exploration is socially, but not privately optimal. Thus, it must be that

$$\frac{c_e}{p_e} - c_i > \beta(V^1_{Estbl} - V^{T+1}_{Estbl}) + K - K\lambda^T \tag{7}$$

However, it is clear that by lowering the cost of the explorative project sufficiently through a research grant, the firm can be induced to explore. If the matching rate is chosen so that the firm is kept indifferent between exploration and exploitation, the firm attains exactly the same payoff as in the firm decision problem.

Proposition 3. *Suppose that Assumptions 1 and 2 hold. Then, the social planner can implement the socially optimal research policy by offering research grants to established firms for explorative projects with matching rates* τ^{RG}_T *that solve*

$$\frac{c_e}{p_e}(1 - \tau^{RG}_T) - c_i = \beta(V^1_{Estbl} - V^{T+1}_{Estbl}) + K - K\lambda^T \tag{8}$$

for T *s.t.* $\bar{T}^* \leq T < \bar{T}$, *where all value functions and* $\bar{T}$ *are obtained from the firm decision problem, and* $\bar{T}^*$ *is obtained from the social planner decision problem. Under such a research grant policy, the firm attains the same payoff as in the firm decision problem.*

Similar to the research grant policy described in Proposition 3, the Austrian Research Promotion Agency favors high-risk projects. The funding policy of the agency depends on factors not directly related to the proposed project, like the "additionality" of the project relative to other projects in the research portfolio of the firm. The research grant policy in Proposition 3 depends on the alternative incremental project that the firm would undertake in the absence of a grant. One implication of this model is that the firm will eventually undertake the explorative project if it is denied public funding, but delay its start. This finding is consistent with survey evidence presented in section 1.5.2 and the perception of the innovation policy expert that grants allow firms to bring inventions ”faster” to the market.

In the Appendix, I compare the research grant policy described in Proposition 3, which selects risky, explorative research projects for funding, with an undirected R&D tax credit policy that uniformly lowers the cost of research for all firms and all projects. I show that a uniform reduction in the cost of research does encourage more explorative research in this setting. Although it is possible to implement the socially optimal research policy with a R&D tax credit, the cost of such a tax credit to the public is substantially higher than the cost of an equivalent research grant policy. The reason is that the tax credit suffers from a double inefficiency in encouraging explorative research: not only are incremental projects subsidized, but also explorative research projects are overfunded. The matching rate of an R&D tax credit that induces the firm to undertake the explorative research project exceeds the matching rate of a research grant that targets the explorative projects. There is a robust and simple intuition: suppose by contradiction that the matching rate of the R&D tax credit is (weakly) lower. By definition, a cost-efficient research grant for the explorative project keeps the firm indifferent between the explorative and the incremental project. Thus, since the tax credit also funds the incremental project, it must be that the firm strictly prefers the incremental project if it is offered a tax credit with a (weakly) lower matching rate instead. In practice, potential savings in public funding from research grant policies need to be compared to the cost of the government agency that administers the research grants. Also, I have not modelled a main impediment to research grants, informational asymmetries.

1.7 Conclusion

How effective are government research grant programs for industrial R&D? This chapter presents evidence from the Austrian Research Promotion Agency FFG of a substantial effect of government research grants on firm patenting outcomes. My estimates suggest that a government research grant increases the propensity to file a patent application with the European Patent Office within 4 years by around 10 percentage points. The average grant size awarded by the agency is 118,000 Euro. This implies that the agency must pay slightly above 1,000,000 Euro in grants in order for one firm to file a European Patent, that otherwise would not have done so. The effect is stronger for established firms above the median age, which is 5 years. I present evidence that research grants encourage established firms to file patents that appear non-trivial with respect to conventionality and number of citations received. Furthermore, receiving a grant is correlated with heavier utilization of technological knowledge novel to the firm.

Consistent with this evidence, I propose an "exploration vs exploitation" model in which the government agency addresses inefficiency in the direction of research. By offering research grants for ambitious high-risk projects, the agency encourages established firms to undertake explorative research on novel research lines, as opposed to incremental research on old research lines. Research grant programs potentially achieve considerable cost savings compared to undirected R&D tax credits in this setting.

My findings contrast with Howell's (2016) findings for the US SBIR program, who finds that government research grants alleviate financial frictions for young firms. There is some room for speculation here. The two programs differ fundamentally in their funding principles. The FFG funds projects with an average matching rate of around 25 percent of total cost. For a young company, a grant that covers 25 percent of the cost may be too low.[32] At an early stage, the firm may be better served with a venture capital deal. On the other hand, even a matching rate as low as 25 percent can have a significant impact on the selection of research projects in established firms. The US SBIR program awards grants of a fixed size and is restricted to small firms. If an applicant firm to the SBIR is old but still small, this may indicate that the firm is not successful. The firm size constraint of the SBIR, which is absent in the FFG, may therefore lead to a bad selection among old applicant firms. Also, the fact that US SBIR grants do not scale to the project cost may work to the detriment of more established firms.

Inefficiency in the direction of research, which may be more relevant for established firms, and financial frictions, which may be more relevant for young firms, constitute seperate hurdles to firm innovation that may require different institutional responses. It is important to study the rules under which subsidies are awarded to gain understanding of their effects.[33] While undirected tax credits may be suitable to raise the aggregate level of R&D spending, they only inadequatly address the aforementioned hurdles. Firms that face financial constraints often have no revenue and thus do not owe corporate income tax. Due to their indiscriminate nature, tax credits inadequatly affect the selection of research projects within firms. Such considerations, among others, suggest that government research grants are potentially important tools of innovation policy. Further research to tease out empirical findings across programs, and related theoretical considerations, are important avenues for further work.

[32] In compliance with EU guidelines, the agency has gradually shifted over time towards discriminating more between small and large firms in its funding policy. However, according to the agency, this was less the case between 2002 and 2005.

[33] With respect to the particular institution studied in this chapter, the Austrian Research Promotion Agency (FFG), it would be interesting to study the effects of grants on the project level, rather than on the firm level. In the course of this project, I matched keywords in the abstracts of the funding applications with the abstracts of firms' patent applications. While my attempt ultimately failed, such an approach may be feasible for certain technical fields or with the help of technical experts.

2 How responsive are inventors to local income taxes? Evidence from residential location choices in Switzerland

2.1 Introduction

Inventors are a key part of the knowledge economy and their locational choices potentially have large consequences for regional development and the emergence of local technology clusters. Policymakers are making concerted efforts to attract highly-skilled, inventive individuals.[34] In this chapter, I study the relationship between personal income tax rates and the residential location choice of inventors in a "border effect" research design, exploiting sharp differences in tax rates across state borders.[35]

The fiscally decentralized structure of Switzerland provides a useful laboratory, as between 38 and 73 percent of the income tax burden of top-income earners is levied locally.[36] In many cases, due to differences in the progressivity of tax systems, geographically proximate municipalities that lie on different sides of state borders exhibit vastly different top income tax rates. As an example, I present a satellite picture (Figure 4) of the two neighboring municipalities Richterswil and Wollerau at the state border between Zürich and Schwyz. Although they lie only 1.5 kilometers apart, the average income tax rate in Wollerau is 14.35 percentage points lower than the average income tax rate in Richterswil for an individual

[34] See the introduction of Moretti and Wilson (2017) for examples of how US states compete for high-skilled workers by offering low income taxes. Three European countries in which highly-skilled foreigners enjoy tax-exemptions are Denmark (studied in Akcigit, Baslandze and Stantcheva (2016)), Netherlands and Sweden. In the case of Denmark, one requirement of the exemption is that the beneficiary undertakes "research work". For an international policy report see OECD (2008) "The Global Competition for Talent: Mobility of the Highly Skilled".

[35] See Black (1999) and Bayer, Ferreira and McMillan (2007) for applications of the border effect research design.

[36] Assumed annual income 500,000 CHF (around 450,000 USD) for unmarried individual in 2010.

earning 500,000 CHF (around 450,000 USD) annually in 2010. I exploit the cross-sectional variation around state borders by focusing on residential location choices in 782 municipalities close to state borders. I estimate a discrete choice model of residential location in which the inventor compares his or her actual observed municipality of residence with a counterfactual neighboring municipality on the other side of the state border. Because of the short distance, many unobserved attributes plausibly remain constant across these two choices. The variation in income tax rates around state borders may be regarded as exogenous if state tax policies are not set according to local political preferences in state border municipalities, which are typically less densely populated and do not include urban centers. I then examine the effect of a state-wide reduction in income tax rates in two states.

Previous studies have shown that inventors migrate in response to changes in income tax rates. Akcigit, Baslandze and Stantcheva (2016) study the location choice of "star" inventors in a panel of eight (mostly) non-neighboring countries and find an elasticity of the number of foreign inventors with respect to the net-of-tax rate (after-tax income) of around 1. Moretti and Wilson (2017) find a corresponding elasticity of 1.7 for inventor migration flows in a panel of US states. While migrating a long distance to a different country or to a different US state entails a profound change in living conditions and high moving costs, there are other settings, like the one analyzed in this paper, in which tax differences may be exploited more easily. In the European Union, there are big differences between the tax systems of densely populated, neighboring countries. With the receding of the barriers to the movement of highly-skilled individuals, a concern is that regions that border low-tax countries may be stunted in their development by "brain drain". This chapter contributes to the unfolding literature on inventor mobility and income tax rates by extending the results of preceding studies to an increasingly local level and by comparing income tax rates to the effect of other amenities that vary at the micro-geographic level.

I find a statistically significant and economically large negative relationship between top income tax rates and the residential location choice of inventors. The elasticity of the number of inventors in a municipality with respect to the net-of-tax rate (after-tax income) is around 4.6. This estimate is considerably higher than the elasticities found by Akcigit, Baslandze and Stantcheva (2016) and Moretti and Wilson (2017). Since neighboring municipalities are close substitutes as residential locations, it seems plausible that inventors are particularly responsive to differences in income tax rates in my setting. Consistent with an explanation whereby domestic inventors have relatively stronger social ties to locations, foreign born inventors exhibit an elasticity with respect to the net-of-tax rate that is 44 percent higher, although this difference is not statistically significant.

As previously mentioned, the research design allows me to study how residential location choices depend on local amenities such as shorter commute, lakefronts and language. Perhaps most interestingly, I find, consistent with Richard Florida's hypothesis that creative types place particular value on "tolerance" (Florida 2002), that inventors appear to have a bias against municipalities where the anti-immigration party SVP ("Schweizerische Volkspartei") secured a larger share of the votes in the federal elections of 2011.

Finally, I document that inventors who have a longer commute to their initial workplace are more likely to transfer to a different workplace over time. In addition, I show that inventors transfer to a workplace that lies closer to their residence. Hence, using personal income tax rates to draw the residences of inventors closer may have positive spillover effects for regions. Low-tax regions that lie within commuting distance of innovation centers are particularly well-positioned to benefit from this dynamic relocation mechanism. This may explain, for example, how the rural low-tax state of Zug, which lies only 30 kilometers south of Zürich, managed to transition from being an under-developed region to an innovation hub while other, more remote, low-tax regions have not seen similar increases in innovative activity.[37]

The rest of the chapter is organized as follows: In section 2.2, I discuss related literature. Section 2.3 describes the data analyzed in this paper. Section 2.4 outlines the empirical strategy. In section 2.5, I present the results, and in section 2.6 I discuss the findings.

2.2 Related literature

Apart from the studies by Akcigit, Baslandze and Stantcheva (2016) and Moretti and Wilson (2017) (discussed in the introduction), the present chapter is related to a number of other studies of tax-induced migration. Young and Varner (2011) find no evidence of a migration response of millionaires to the increase in New Jersey's personal top income tax rate. Bakija and Slemrod (2004) find evidence of modest migration of high-income earners in response to changes in state income taxes in a panel of US states. Kleven, Landais and Saez (2013) find that European Soccer players are highly responsive to top tax rates.[38] The present chapter is not the first to exploit the fiscally decentralized structure of Switzerland to study the effect of income tax rates: Kirchgässner and Pommerehne (1996) find that low-tax states exhibit higher shares of high-income earners. Schmidheiny (2006) finds that high-income households are more likely to locate in low-tax municipalities within the Basel metropolitan area. Liebig, Puhani and Sousa-Poza (2006) find that changes in tax rates in municipalities between 1995 and 2000 predict out-migration of university graduates.

[37]See, for example, "Low tax Zug aims to become Switzerlands Crypto Valley", Reuters 10/08/2016 (linked), "How Switzerland became the Silicon Valley of Robotics", Forbes 10/26/2017 (linked) or "Welcome to Zug: the sleepy Swiss town that became a global hub", The Guardian 05/30/2008 (linked)

[38]See also Cohen, Lai and Steindl (2011).

Furthermore, this chapter is related to a strand of the literature that studies the locational preferences of inventors and scientists that are non-tax related. Miguélez and Moreno (2013) find that inventor migration flows between European regions are predicted by physical proximity and cultural, social and institutional similarities. Dorner, Harhoff, Hinz, Hoisl and Bender (2017) find that Western regions in Germany with stronger historically determined social ties to Eastern regions attracted more inventors from East-Germany after the fall of the iron curtain. Dahl and Sorensen (2010) study the choice of Danish scientists and engineers of where to work and find that they do not exploit regional differentials in wages because of social ties to partners, family and former classmates.[39] This chapter is also related to a literature that sprouted after Richard Florida's influential work on the "creative class" (see Florida 2002), which posits that creative individuals are attracted by the prevalence of "tolerance" at a location. Although preliminary evidence is consistent with the hypothesis, researchers have raised questions about the interpretation of these results.[40]

Last, the chapter is methodologically related to a large number of studies that apply McFadden's multinomial logit model to problems of location or product choice. Multinomial logit models similar to the one used in this chapter have been used, for example, by Kleven, Landais and Saez (2013) and Akcigit, Baslandze and Stantcheva (2016) to study migration responses to tax rates.

[39] Other studies of the locational choices of highly skilled include Gottlieb and Joseph (2006), Scott (2010), Faggian and McCann (2009), Brown and Scott (2012) and Dorfman, Partridge, and Galloway (2011).

[40] See studies by Bereitschaft and Cammack (2015), Wojan, Lambert and McGranahan (2007), Frenkel, Bendit and Kaplan (2013) and Lawton, Murphy and Redmond (2013). Glaeser (2005) criticizes that the impact of "tolerance" cannot be disentangled from other correlated factors that may explain the agglomeration of highly-educated, "creative" individuals, such as quality of education and availability of top universities.

2.3 Data

The three sources of data used in this chapter are inventor data from the patent database by Miguélez and Fink (2013), data on income tax rates from the Swiss Federal Tax Administration and other municipality-level data from the Swiss Statistical Office.

2.3.1 Inventors

The inventor dataset is based on the inventor mobility database by Miguélez and Fink (2013). They extract information from patent applications filed under the Patent Cooperation Treaty (PCT).[41] The crucial feature of the patent applications included in this database is that patent applicants are required to declare the nationality of the inventors listed on the patents. This feature is the primary reason for using this data.[42] A secondary reason for using this data is that PCT patents capture the commercially most valuable patents.

Miguélez and Fink (2013) combine information on the inventor's nationality and country of residence to measure aggregate international inventor migration flows. I extract all inventors who reside in Switzerland and who were listed on at least one PCT patent filed between 2005 and 2012.[43] The patent contains the address of the inventor as of the date of the application. In addition, I collect address data for the applicants listed on the PCT patents via Patentscope, the bibliographic database of the World Intellectual Property Organization. Since my aim is to match inventors to their presumed employers, I only retain

[41]The relationship between income tax rates and number of PCT inventors at the country-level is also analyzed in Akcigit, Baslandze and Stantcheva (2016).

[42]Typically, the inventors' nationalities are not known and can only be inferred (rather imperfectly) by origin of name. However, identification of inventor migration flows within Europe is very unreliable with this method. In the case of Switzerland, which is linguistically fragmented into German-, French- and Italian-speaking parts (and a Raetheromanian-speaking part to be complete), the dataset by Miguélez and Fink (2013) enables me to identify German, French and Italian inventors that live in Switzerland, who would otherwise be indistinguishable from Swiss nationals.

[43]The year 2012 is the last year for which data is available in the database of Miguélez and Fink (2013) .

inventors that can be related to a unique patent applicant with an address in Switzerland.[44] I geocode the residential address of the inventor and the address of the presumed employer at the municipality level and disambiguate inventors by name. Henceforth, I will refer to the municipality where the presumed employer is located as the "workplace" of the inventor, although in reality there may be cases where the stated address of the presumed employer deviates from the place where the inventor performs research.

In total, there are 12,787 distinct inventors, of which 5,016 are foreign-born,[45] and 14,948 distinct inventor-residence-workplace triples in my dataset. If an inventor appears in more than one municipality of residence or in more than one workplace, i.e. if the inventor moves or switches workplace, the inventor is included more than once in the dataset.[46] However, depending on the context, I will refer to an inventor-residence-workplace triple simply as an inventor for reasons of legibility. Figure 5 shows the spatial distribution of inventor workplaces in my dataset. Many inventors are employed in the urban centers Zürich, Basel, Geneva and Lausanne. Figure 6 shows the spatial distribution of inventor residences. The distribution of the distance between the municipality of residence and the municipality where the workplace is located, which I refer to as the "commuting distance", is presented in Figure 7. 23.4 percent of inventors live in the same municipality as they work in, 40.9 percent have a strictly positive but short commute of less than 21 km (13.125 miles), 18.5 percent commute between 21 km and 51 km (31.875 miles) and 7.8 percent live between between 51 km and 81 km (50.625 miles) away. All distances are "straight-line" distances between the

[44] Around 65 percent of all PCT patents that were considered had exactly one applicant with a Swiss address listed.

[45] I regard inventors, who are disambiguated by name, as foreign-born if they appear with a nationality other than Swiss on at least one patent application. Some inventors acquire Swiss nationality over time. The nationalities of the foreign-born inventors are as follows: the largest groups are Germans (40 percent) and French (14 percent), followed by Italians (9 percent), British (6 percent) and Austrians (3 percent). Another 10 percent stem from other EFTA nations (EU or Norway), 6 percent from either the US, New Zealand, Australia or Canada and 12 percent from the rest of the world (most notably Japan, China and India).

[46] For inference, I cluster all bootstraps at the inventor level.

centroids of the municipalities, not taking into account the actual topology of road ways or public transport.

2.3.2 Income tax rates

In Switzerland, income taxes are levied by the 26 states ("cantons"), the municipalities and by the federal republic. Each state has its own tax system and legislation that stipulates how tax authority is shared between the state and the municipalities. I collect harmonized data on local income tax rates for municipalities in Switzerland from the Federal Tax Administration. The income tax rates that I use throughout this chapter are given by the tax burden in percent of the individual gross income levied by the state and the municipality, excluding the federal income tax, for a single income earner with no children for different income levels.[47] The tax base that is subject to the income tax is very broad and encompasses, among other sources, wage income, capital income, rental income and the imputed rental value of owned real estate. Income taxes are solely based on the location of residence. In the main analysis, I use the tax rate for the year 2010, which is the first year for which data on all municipalities is available. Before 2010, coverage is spotty and varies by state. Since inventor address data stems from the years 2005 to 2012, there is a discrepancy between the time when the tax rate is measured and the time when the inventor location is measured. However, I expect this inaccuracy to have an attenuating effect on my estimates.

[47] Foreign nationals without permanent residence permit in Switzerland, that earn wage income not in excess of 120,000 CHF are taxed according to a different (but highly correlated) income tax schedule. This income tax schedule depends only on the state of residence, but not on the municipality. Foreign nationals from EFTA (EU or Norway) member states, with exceptions for Bulgaria and Romania, obtain a permanent residence permit after 5 years. Rules for other foreigners vary, but typically they are eligible to apply for a permanent residence permit after 10 years. I assume that the normal income tax schedule is the relevant tax schedule for the foreign-born inventors in my sample. First, it seems likely that they earn wage income in excess of 120,000 CHF. Second, they may have obtained a permanent residence permit by the time that they appear as inventors on a patent application. Third, even if they have not obtained the permit yet, they may have chosen their residence in a forward-looking way.

Similar to Moretti and Wilson (2017) and Akcigit, Baslandze and Stantcheva (2016), I do not observe actual inventor income. Following the aforementioned papers, I focus on tax rates for the top income percentiles. Since the inventors in my dataset are inventors of PCT patents, which capture the commercially most valuable patents, this assumption seems sensible.[48] I assume an annual income in the range from 200,000 CHF (around 180,000 USD in 2010) to 500,000 CHF (around 450,000 USD in 2010), thereby covering the range from the 97th to the 99.5th percentile in taxable income. I will refer to the income tax rate for an annual income of 500,000 CHF as the “top income tax rate”. My main results are presented for an annual income of 500,000 CHF, but I consider annual income levels of 200,000 CHF, 300,000 CHF, and 400,000 CHF in the robustness section. To be clear, all tax rates are average tax rates, not marginal tax rates. Figure 8 shows a map of top income tax rates in Switzerland. Corresponding maps for annual income levels of 200,000 CHF, 300,000 CHF and 400,000 CHF can be found in the Appendix. Because some states have highly progressive tax systems while others have “flat” tax systems, the range of local income tax rates grows with the level of income. As a result, there are massive differences in income tax rates for high and top income tax brackets. The top income tax rate is only 6.67 percent in Wollerau (Schwyz), but 28.62 in Vermes (Jura).

The variation in top income tax rates is mostly driven by the state component. This results in a marked “state border effect”, evident in Figure 8. In the introduction, I mentioned the state border region between Zürich and Schwyz along the southern shore of Lake Zürich as an example. A satellite picture of the border region is presented in Figure 9. The municipalities of Richterswil (21.02%) and Wädenswil (21.59%) in the state of Zürich have considerably

[48] Akcigit, Baslandze and Stantcheva (2016) use citation data and the total number of patents produced by an inventor to identify successful “star inventors”. They discuss evidence from the US, Finland and Germany that inventor quality is strongly linked to inventor salary. Based on US IRS data from Bell et al. (2015), they suggest 200,000+1,400*(number of citations) as the best quality-adjusted predictor of annual inventor salary.

higher top income tax rates than the municipalities of Wollerau (6.67%), Feusisberg(6.95%) and Freienbach (6.89%) which lie in the state of Schwyz. By relocating only 1.5 km from Richterswil to Wollerau, an individual earning 500,000 CHF annually can save 71,750 CHF in taxes. However, at the same time, Richterswil and Wädenswil may offer other municipal amenities that inventors might care about. I will return to this example in the next section.

2.3.3 Other data on municipalities

Data on municipal amenities that may influence the location choice of inventors is obtained from the Swiss Statistical Office for the year 2010. Summary statistics for all variables can be found in Table 9. The selection of the variables is guided by the literature in Urban Economics.

For each municipality, I record school graduation rate,[49] crime rate, municipality population size, public transport usage, the distance to the national border and whether or not the municipality has a lakefront. More information on these variables can be found in the Appendix.

School graduation rate, public transport usage and crime rate are potentially related to local top income tax rates, since education, infrastructure and public safety are partially or wholly funded by states and/or municipalities. In my main results, I include these variables

[49]It is given by the share of 19-year olds who complete school with an advanced degree that is a requisite for university studies called “Gymnasiale Maturität”. The attainment rate ranges from 7.14 percent in rural Kulm (Aargau) to 42 percent in Lavaux-Oron (Waadt) near Lausanne. Schools are administered by states and students are allocated to public schools based on the municipality of residence. Parents may decide to opt-out of public schooling and instead enroll in a private school. Private schools are not subject to residence allocation rules and account for around 13 percent of the total number schools, evenly distributed across states. Due to the institutional complexity and because public authorities are actively trying to deter strategic sorting into schools, it is notoriously difficult to find data on school quality. Since access to universities is regulated federally, I expect the academic standards for attaining “Gymnasiale Maturität” to be similar across all states. The “Gymnasiale Maturität” is a rather exclusive degree whose attainment entails careful selection of schools from an early age on. In 2010 only, 20 percent of 19 year olds in the general population attained “Gymnasiale Maturität”. See, for example, Black (1999) and Bayer, Ferreira and McMillan (2007) for evidence on location choices and school quality for the US.

as controls. I assume in the interpretation of the relationship between the top income tax rate and the location choice that the policymaker has enough flexibility in the budget to hold, for example, spending on schools constant when changing the tax rate.[50]

Furthermore, for each inventor-municipality pair, I compute a synthetic measure of commuting distance, which is given by the "straight-line" distance between the centroid of the municipality where the workplace of the inventor is located and the centroid of the municipality where the inventor resides (or may consider to reside). It may deviate from the actual travel distance to the place where the inventor performs research for two reasons: first, it does not take into account the actual topology of roadways or public transport and is likely to understate the travel distance by car or public transport, especially for mountainous regions. Second, as mentioned in section 2.3.1, there is some inaccuracy in asserting the workplace of the inventor. In particular, for inventors that reside far away from their presumed employer, I cannot tell whether they work from home or commute infrequently or whether their workplace is simply mismeasured. To address the first concern, I exclude mountainous regions from the sample as a robustness check. To address the second concern, I also present results conditional on the inventor residing no further than 80 km away from the presumed employer.

In addition, I collect data on the percentage of German speakers and French speakers (as

[50]In my research design, differences in roads and transportation should be negligible for neighboring municipalities. Municipalities typically fund roads, while states subsidize or own private bus and train networks. However, the overwhelming amount of funding for infrastructure comes from federal sources, which also finances the Federal Swiss railway company SBB. Other important state spending categories are social transfers and health. Since inventors are typically high-income earners, I assume that social transfers, which are redistributive, do not matter for their choice of residence. All states are required to offer non-discriminatory insurance (through private health insurers) to all state residents. Premiums and health services are regulated federally, but differ slightly by state. The basic insurance typically only covers treatment in the state of residence. If a state lacks a particular medical facility (e.g. a hospital with an oncology department), it designates contracted out-of-state facilities. Average spending for health insurance premiums (in the state) ranges between around 4100 CHF annually in Appenzell Inerrhoden and 6800 CHF in Basel Stadt annually in 2017 (www.bfs.admin.ch "Krankenkasse: Prämienregionen 2017").

primary language) in a municipality and the vote share of the anti-immigration party SVP ("Schweizerische Volkspartei") in the federal elections of 2011. Because a sizable number of foreign-born inventors stem from France or Germany, I can study the match of the primary language spoken by the inventor and the primary language spoken in the municipality.

Returning to the example of the state border region between Zürich and Schwyz, discussed in the previous section, the satellite picture in Figure 9 shows that the muncipalities of Wädenswil and Richterswil lie closer to the city of Zürich. Furthermore, in contrast to the municipality of Feusisberg, they are situated right at the lake. The municipalities on the two sides of the state border differ in other ways: for example, the schools in the municipalities of Wollerau, Feusisberg and Freienbach exhibit a slightly higher graduation rate (24.18 percent compared to 21.76 percent). The vote share of the SVP was on average 10 percentage points higher in these three municipalities than in the two municipalities on the other side of the state border (40.1 percent compared to 30.7 percent).

2.4 Empirical Strategy

In 4.1, I provide a discussion of the research design and the main assumptions that underlie the identification strategy. In section 2.4.1, I present preliminary reduced-form evidence and summary statistics for the relevant population studied in this paper. In section 2.4.2, I introduce the empirical model of residential location choice and discuss in detail on how it is estimated.

2.4.1 Preview and identifying assumptions

I estimate a model of residential location choice using an approach akin to the "border effect" research design that was introduced in the seminal contribution of Black (1999). It has two key features: first, in order to control for unobserved attributes that are spatially correlated, I limit choices for inventors to comparisons between municipalities that lie no farther than 5 kilometers (3.1 miles) apart.[51] Second, I utilize the state border as a source of sharp variation in top income tax rates. Although top income tax rates vary at the municipality level also within states, the variation is mostly driven by the state component. This results in a marked difference in top income tax rates between municipalities that lie on different sides of the state border, evident in Figure 8. I construct the instrumented income tax rate in a municipality as the average of the income tax rates of all other municipalities in the same state that lie farther than 5 kilometers away. Using the instrumented income tax rate instead of the actual income tax rate in a municipality addresses concerns that the municipality-component of the local income tax rate is endogenous to local political preferences. Differences in the municipality-component of the local income tax rate between neighboring municipalities may be explained by differences in local political preferences.

A crucial question is whether differences in local income tax rates that can be attributed to state tax policies can be regarded as exogenous for neighboring municipalities. In my setting, it seems reasonable to assume that average income tax rates at the state level are not set according to the local political preferences in specific state border municipalities, which are typically less densely populated and do not include urban centers. Only two of the 25 largest cities, Sankt Gallen and Schaffhausen, are included among the state border municipalities.[52] The average income tax rates in the state may reflect spending on public infrastructure that is concentrated in larger cities, like spending on universities and hospitals. Furthermore, state legislators are less likely to cater to state border regions because of

[51] A major concern is that the counterfactual income that an inventor can attain at a particular location depends on the set of employers that lie nearby. Since demand for inventors is geographically concentrated, municipalities across Switzerland differ vastly in this respect.

[52] See definition further below. The city of Basel, which is the only city that is also a state, and the surrounding state "Basel Land" were collapsed to a single state in the analysis.

the relatively lower number of voters. Consider the example of the border region between the states of Zürich and Schwyz, discussed in section 2.3. The fact that the top income tax rates on the Zürich-side of the border are considerably higher than in the neighboring municipalities on the Schwyz-side of the border, may be the unintended consequence of state level taxes being set according to the preferences of the urban city center of Zürich. In the robustness section, I show that my findings are robust to excluding small states where local political preferences in state border municipalities have a larger influence on state tax policies.

However, concerns remain that the municipalities on different sides of state borders may differ, even if they lie no farther than 5 km apart. As a first step, I include controls for municipal amenities which are deemed relevant in Urban Economics. Furthermore, I examine whether reductions in income tax rates in two states lead to observed changes in choices between municipalities on different sides of state borders over time. Unobserved differences between municipalities that are persistent or do not change fast, like culture or political attitudes, are less likely to bias my estimates in this setting.

Preliminary reduced-form evidence and summary statistics

This section presents reduced-form correlations and summary statistics for the relevant population studied in this paper. In my research design, I focus on the residential location choice of inventors between geographically proximate municipalities around state borders. Consequently, the population of inventors studied in this chapter are inventors who live close by state borders. I will discuss this limitation in the context of the residential location choice model in section 2.4.2. I classify a municipality as a "state border municipality" if it lies within 5 kilometers of distance of a municipality that belongs to a different state. Furthermore, I eliminate a subset of municipalities that are neither the residence of any inventor, nor lie close to any inventor's municipality of residence, since the observed residential location choices are not revealing for such municipalities. I will define these sets more carefully in section 2.4.2, when I introduce the residential location choice model.

In Figure 10, I show the relationship between the difference in the (actual) top income tax rate and the difference in the number of resident inventors for all distinct pairs of municipalities that lie on different sides of state borders and no farther than 5 kilometers apart. The negative relationship suggests that a 5 percentage point lower top income tax rate, relative to the municipality on the other side of the state border, is associated with 2.5 additional inventors in a municipality. The relationship is robust to normalizing the number of inventors by the total population in a municipality.

Table 9 presents summary statistics for municipalities and Table 10 shows summary statistics of the characteristics of the municipalities of residence at the inventor-residence-workplace level. 782 out of 2580 municipalities are included in the set of state border municipalities. Compared to the set of all municipalities, state border municipalities are on average smaller (2622.54 compared to 3048.86 in average population size) and are home to fewer inventors (4.26 compared to 5.45 inventors per municipality). Relatedly, when I rank inventors by the population size of the municipality of residence, I find that the median population size is smaller for inventors who live in state border municipalities than for all inventors (5,468 compared to 10,812). 3,549 out of 14,948 inventors reside in state border municipalities[53]. The average commuting distance for inventors who live in state border municipalities is similar to the average commuting distance for all inventors (28.01 km compared to 28.03 km).

[53]There are 3549 distinct inventor-residence-workplace triples. Since the characteristics of the municipality of residence change when the inventor moves or switches workplace (the commuting distance changes), all statistics are at the inventor-residence-workplace level, not at the inventor level.

2.4.2 The residential location choice model

I will now turn to the residential location choice model. I assume that the utility of inventor $i \in I$ living in municipality $m \in S$ is given by the multinomial logit model

$$u_{im} = \beta_1 f(\tau_{im}, Y_{im}) + \beta_2 X_m + \beta_3 Z_{im} + \epsilon_{im} \tag{9}$$

where τ_m is the income tax rate in municipality m, which is either the instrumented[54] or the actual tax rate, Y_{im} is the before-tax income of the inventor in municipality m, X_m is a vector of municipality-specific covariates and Z_{im} are covariates specific to inventor-municipality pairs. ϵ_{im} is an i.i.d. error term that is assumed to be extreme-value distributed. The inventor is assumed to choose the municipality that promises maximal utility.

For any inventor consider the set of municipalities within 5 kilometers of distance to the inventor's municipality of residence. When the inventor compares his or her actual municipality of residence with any other municipality within 5 kilometers of distance, many factors that play a role in the inventor's residential location choice should be the same or very similar across all locations. Most importantly, it is reasonable to assume that the inventor can attain the same before-tax income at all locations. Any employer that lies within a reasonable commuting distance of the inventor's municipality of residence also lies within a reasonable commuting distance of the other municipalities. Differences in climate, which were found to be an important factor for location choices in the United States (see Glaeser and Tobio 2008 , Albouy 2008, 2016, Albouy et al. 2013) are negligible. Apart from income and climate, I expect the value derived from infrastructure and amenities that are only used infrequently to be similar, like hospitals, hiking trails or venues for cultural events. This

[54] The instrumented income tax rate in a municipality is given by the average of the income tax rates of all other municipalities in the same state that lie farther than 5 km away.

holds true because small differences in travel distance should be inessential for rare events like hospital visits. Furthermore, because of the short distance between the municipalities, the costs of many goods and services are similar.

As discussed, the second part of the empirical strategy relies on the state border as a source of sharp variation in top income tax rates. Consider the set of municipalities within 5 kilometers of distance to the inventor's municipality of residence that belong to a different state, and the inventor's municipality of residence itself. While attainable before-tax income is constant in this set, after-tax income varies by construction. Since the inventor's municipality of residence belongs to a different state than the other municipalities, there is potentially large variation in income tax rates despite the geographical proximity.

To absorb some of the variation between municipalities, I include controls for the amenities listed in section 2.3.3. Furthermore, I posit that the impact of a subset of amenities, which are exogenous scenic or geographic attributes, on the residential location choice of inventors is identified within the framework of my research design. I argued that small differences in travel distance to amenities that are only accessed infrequently may be inessential in the choice of location; on the other hand, small differences in the distance of the daily commute are still important. For this reason, this research design is conducive to assessing the importance of commuting distance, although I cannot rule out that it is correlated with other attributes that depend on the proximity to the workplace. Similar considerations apply for the distance from a municipality to the national border, which may be particularly relevant for inventors that stem from neighboring countries. Furthermore, by comparing municipalities that possess a lakefront with geographically proximate municipalities that do not, I distinguish its effect from other scenic attributes that vary continuously.

In preparation for the description of the estimation strategy further below, it is necessary to introduce some notation. The set of municipalities that includes the municipality of residence of inventor i, m_i, and all municipalities within 5 kilometers of distance that lie in a different state (than m_i) is defined as

$$M_{5km}(m_i) = \{m_i\} \cup \{m' \in S | d(m_i, m') < 5 \text{ km } \wedge \text{ state}(m_i) \neq \text{state}(m')\}$$

where S is the set of all municipalities in Switzerland and d is the distance between two municipalities.[55] In Figure 11, I illustrate the set of municipalities in $M_{5km}(m_i)$ for an inventor living in Richterswil (Zürich) at the state border between Zürich and Schwyz, which includes Wollerau (Schwyz), Feusisberg (Schwyz) and Richterswil (Zürich) itself. The collection of sets $\left(M_{5km}(m_i)\right)_{i \in I}$ will be used to construct the inventors' choice sets. Consider the set of municipalities that lie close by a state border and that are either the residence of some inventor or that lie within 5 kilometers of distance to some inventor's municipality of residence. This set is defined as

$$SB_{5km} = \left\{ \bigcup_{i \in I} M_{5km}(m_i) \middle| M_{5km}(m_i) \setminus \{m_i\} \text{ is non-empty} \right\}$$

The set of municipalities SB_{5km} is the set of state border municipalities referred to in section 2.4.1.[56] In the multinomial logit model described by (1) (and for MNL models more generally), if the inventors' preferences are homogeneous, then all parameters of the model can be estimated from any fixed subset of alternatives. In particular they can be estimated from the subset SB_{5km}. In this case, I consistently estimate the parameters of the model

[55] The distance between two municipalities is given by the "straight-line" distance between the centroids of the municipalities.

[56] It is worth pointing out that a municipality that is actually located at the border of a state but has a very large surface area, may be dropped from the set SB_{5km} simply because there may be no other municipality within 5 kilometers of distance of the centroid of the municipality. Therefore, I also consider alternative ways of defining the choice sets in the robustness section.

from the subset of inventors who reside close by state borders. However, if preferences are heterogeneous, then the question arises whether inventors living close by state borders and inventors living far from the state border have the same preferences. This consideration applies especially to inventors who live in the urban centers, which do not lie at state borders.[57] The estimates obtained with this approach may therefore only apply to the population of inventors who reside close by state borders, but not to all inventors. However, in the rest of the chapter I assume that preferences are homogeneous and revisit the issue of preference heterogeneity in the robustness section. Summary statistics for state border municipalities are discussed in section 2.4.1.

Estimation of the residential location choice model

The empirical strategy described in the previous section is implemented as follows: for any inventor i, living in municipality m_i, a pair of municipalities from $M_{5km}(m_i)$ is assigned randomly to serve as the choice set. One element is i's actual municipality of residence m_i, while the other element of the pair is a counterfactual municipality drawn from $M_{5km}(m_i) \setminus \{m_i\}$, i.e. a municipality that lies no farther than 5 kilometers away from the municipality of residence and on the other side of the state border. The distribution over choice pairs for inventor i is given by

$$P\Big(\{m_i, m'\}|i\Big) = \frac{1}{|M_{5km}(m_i)| - 1} \text{ for all } m' \neq m_i \text{ in } M_{5km}(m_i) \tag{10}$$

Consider any particular random assignment of choice pairs. Let m'_i be the counterfactual municipality in inventor i's choice pair that i does not reside in. Then, for any simulated

[57] The city of Basel, which is the only city that is also a state, and the surrounding state "Basel Land" were collapsed to a single state in the analysis.

random assignment of choice pairs, I compute the modified likelihood function

$$L(\beta) = \sum_{\{i \in I | m_i \in SB_{5km}\}} \log \left\{ \frac{e^{U_i(m_i,\beta) - \log(|M_{5km}(m_i)|-1)}}{e^{U_i(m_i,\beta) - \log(|M_{5km}(m_i)|-1)} + e^{U_i(m_i',\beta) - \log(|M_{5km}(m_i')|-1)}} \right\} \tag{11}$$

for the model described by (1), where

$$U_i(m,\beta) = \beta_1 f(\tau_{im}, Y_{im}) + \beta_2 X_m + \beta_3 Z_{im} \tag{12}$$

The maximizer of the modified likelihood function L is a consistent estimator of β^*, as was shown in the seminal work by McFadden (1978), who laid out in detail how multinomial logit models may be estimated from random subsets of alternatives.[58] All attributes that are constant across all municipalites in $M_{5km}(m_i)$ drop out and are not considered in the estimation.[59] I consider two different specifications for the function f of the tax rate and before-tax income. In the the first specification, I follow Akcigit, Baslandze and Stantcheva (2016) and Moretti and Wilson (2017) and let $f(\tau_{im}, Y_{im})$ be given by $\log((1-\tau_{im}/100)Y_{im}) = \log(1-\tau_{im}/100) + \log(Y_{im})$. The coefficient of $\log(1-\tau_{im}/100)$

[58]The formal statement is as follows:

Corollary 1. *(McFadden 1978, Ch.7, Theorem 2) For any simulated random assignment of choice pairs, let $L(\beta)$ be the associated modified likelihood function. Then, $\hat{\beta} = argmax_\beta L(\beta)$ converges in probability to β^* as the number of inventors $|I| \to \infty$.*

Proof. Note that the distribution P, from which the choice pairs are drawn, does satisfy the "Positive Conditioning Property" in McFadden (1978). The property states that for any pair of alternatives (a,b), choice set D and distribution P, it holds true that $b \in D$ and $P(D|a \text{ was chosen}) > 0$ implies $P(D|b \text{ was chosen}) > 0$. This is satisfied in my case because for any revealed choices by inventors i and j, m_i and m_j, $m_j \in M^{SB}_{5km}(m_i)$ implies that $m_i \in M^{SB}_{5km}(m_j)$ (since distance is symmetric) and hence

$$P\Big(\{m_i, m_j\}|i\Big) > 0 \text{ implies } P\Big(\{m_i, m_j\}|j\Big) > 0$$

Therefore, I can apply Theorem 2 in McFadden (1978) and use the modified likelihood function specified in the Theorem. □

[59]Note also that this estimation strategy significantly relaxes the IIA assumption implicit in MNL model. I assume constant cross-tax-elasticities only between pairs of neighboring municipalities, not on the global set of municipalities.

is then interpreted as the elasticity with respect to the net-of-tax rate. In this specification, I account for other taxes by adding the federal income tax rate (which is constant for all municipalities) to the income tax rate τ_{im}. In the second specification, I let $f(\tau_{im}, Y_{im})$ be linear. The advantage of the first specification is that the coefficients are comparable to the results of Akcigit, Baslandze and Stantcheva (2016) and Moretti and Wilson (2017). The advantage of the second specification is that it allows me to compare the coefficient of the tax rate with the coefficients of other municipal amenities. Because I assume that counterfactual before-tax income is constant across municipalities in $M_{5km}(m_i)$ (explained in section 2.4.2), before-tax income drops out of the modified likelihood function in both specifications, and is consequently not considered in the estimation.

The estimated coefficients are interpreted as the effect of a one-unit change in a variable on the logarithm of the odd-ratio (as in any multinomial choice model), or in the case of a logarithmized variable, as the effect of a one-percent change on the logarithm of the odd-ratio. Following Akcigit, Baslandze and Stantcheva (2016) and Moretti and Wilson (2017), the average (semi-)elasticity of the probability to locate in a municipality with respect to a variable is then given by

$$E\left[\frac{d \log P}{dX}\right] = \beta E[1 - P]$$

Because the number of locations is large in my model, the term $E[1 - P]$ is very close to 1 and the average (semi-)elasticity is therefore approximately equal to the coefficient β.

In principle, the random draw of choice pairs can be simulated arbitrarily often to obtain different estimators $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$,.. and so on. I compute the consistent estimator

$$\hat{\beta}^T = \frac{1}{T} \sum_{t=1}^{T} \hat{\beta}_t \tag{13}$$

where T is the number of simulations. Intuitively, I hope to mitigate the uncertainty that solely stems from the simulation of the choice pairs with this strategy. I estimate confidence intervals for $\hat{\beta}^T$ by re-sampling inventors with replacement and obtain a set of estimates for $\hat{\beta}^T$.[60] To keep this procedure computationally feasible, I choose $T = 100$ and re-sample inventors 500 times. As a goodness-of-fit measure, I compute the average McFadden R-squared $\frac{1}{T}\sum_{t=1}^{T}(1 - L_t/L_0)$, where L_t is the value of the modified likelihood evaluated at the maximizer $\hat{\beta}_t$ and L_0 is the modified likelihood at $\beta = 0$. The reported number of observations is the sum of the number of alternatives per inventor $\sum_{\{i \in I | m_i \in SB_{5km}\}} |M_{5km}(m_i)|$.

On a conceptual level, it is important to note that the estimates obtained with this approach are not structural "willingness-to-pay" estimates, as for example in Bayer et al. (2007).[61] For example, a lower tax rate may attract high-income earners from many different professions. To the extent that inventors prefer having "rich neighbors", this indirect effect is also reflected in the estimated coefficient of the income tax rate.

2.5 Results

This section presents the results for the residential location choice model described in the previous section. Section 2.5.1 documents a strong relationship between the choice of residence of inventors and the top income tax rate. I contrast domestic and foreign-born inventors and present the results of selected robustness checks. Section 2.5.2 presents supporting evidence from changes in income tax rates in two states. Results for other amenities in the cross-section are shown in section 2.5.3.

[60] Re-sampling is carried out at the inventor level, not at the inventor-residence-workplace level.

[61] See Bayer, Ferreira and McMillan (2007) for a discussion of how sorting and endogenous rental prices indirectly affect "the value of amenities".

2.5.1 Evidence on top income tax rates

In Table 11, I present the estimates of the elasticity of the probability to locate in a municipality with respect to the net-of-tax rate. In columns 1-3, I present the results for the main specification that uses the instrumented top income tax rate, which is the average of the top income tax rates of all other municipalities in the same state that lie farther than 5 kilometers away.[62] I find an elasticity of 5.94 in the parsimonious model that does not include controls (column 1). When I include all control variables, the estimate is revised to 4.57 (column 2), which I regard as my main result. Compared to the results of Akcigit, Baslandze and Stantcheva (2016) and Moretti and Wilson (2017), who find corresponding elasticities of 1 and 1.7 respectively, top income tax rates appear to have large effects in my sample. The strong association between the residential location choice of inventors and top income tax rates suggests that attenuating factors, like potentially lower supply of public goods in low-tax municipalities (which are not fully accounted for by my control variables) or higher housing prices, do not compensate for the difference in the tax burden of inventors.[63] A mechanism that may have a magnifying effect is that differences in income tax rates induce sorting by income, and that inventors may have a preference for having high-income earners in their immediate neighborhood.

[62]The advantage of using the instrumented top income tax rate is that the identified effect is solely due to differences in the average level of taxation between the states. If state tax policies are not set according to local political preferences in state border municipalities, this variation may be regarded as exogenous. The municipal component of the local tax may be endogenous to local political preferences in state border municipalities. The advantage of using the actual top income tax rate is that the actual top income tax rate is an *exact* measure of the local tax burden.

[63] The main public spending categories of states and municipalities are (in order of spending amounts) education, social transfers, health, public safety and transport/roads/sanitation (data from www.bfs.admin.ch, "Ausgaben nach Funktionen, Kantone, FS-Model"). In the main specification, I have included controls for school graduation rate, public transport usage and crime rate. I assume that social transfers are not relevant for inventors, since they enjoy high incomes. There are also small differences between the public health systems in states, discussed in section 2.3 footnote 50. Other public spending by states and municipalities include environmental protection, culture and general administration. It is also important to keep in mind that the differences in top income tax rates are to a substantial extent due to differences in the *progressivity* of tax systems between states, and that differences in tax rates for average incomes are generally smaller. Finally, there are federally regulated inter-state transfers to reduce budget imbalances.

This result suggests, for example, that having a top income tax rate of 15 percent instead of a top income tax rate of 20 percent is associated with an increase in the expected number of inventors in the municipality of approximately 33 percent. As a thought experiment, consider two neighboring municipalities that are identical in every way, except that one municipality has a top income tax rate of 15 percent, while the other one has a top income tax rate of 20 percent, and consider the relative choice between only these two municipalities: the difference in the top income tax rates leads to a shift in the relative choice probabilities for any inventor from 50:50 to approximately 57:43 in favor of the low-tax municipality.

The main estimate of the elasticity with respect to the net-of-tax rate is robust to including quadratic terms for the control variables commuting distance and the distance to the national border and cubic terms for municipality population size (column 3). In columns 4-6, I show the results when I use the actual top income tax rate in the municipality instead of the instrumented top income tax rate. In this specification, the estimate of the elasticity of the probability to locate in a municipality with respect to the net-of-tax rate is around 5.5. The elasticity of the probability to locate in a municipality with respect to the net-of-tax rate is statistically significant at the 0.01 level in all models.

Foreign-born vs domestic inventors

I present the results when I split the sample between foreign-born and domestic inventors in Table 12. In the main specification that uses the instrumented top income tax rate, I find that the estimated elasticity of the probability to locate in a municipality with respect to the net-of-tax rate is around 44 percent higher for foreign-born inventors than for domestic inventors (columns 1 and 2). This observation is consistent with the notion that domestic inventors have stronger idiosyncratic social ties to locations. To test the hypothesis that foreign-born inventors are more responsive than domestic inventors, I also estimate a model with an interaction-term using the pooled sample of inventors (column 3). In a one-sided test, I cannot reject the hypothesis that the elasticity is higher for domestic inventors ($p = 0.15$). When I use the actual top income tax rate, the difference in the responsiveness between domestic and foreign-born inventors appears smaller and is only around 20 percent (columns 4 and 5).

The higher goodness-of-fit of the model for foreign-born inventors suggests that the model is better in predicting the residential location choices of foreign-born inventors than the residential location choices of domestic inventors.

Robustness checks

Because of the substantial computational cost of obtaining confidence intervals (see section 2.4.2 for how they are obtained), I only present point estimates in this section. Unless otherwise stated, I discuss the estimates from specifications that use the instrumented income tax rate[64] and that include linear controls for all local amenities (as in Table 11 column 2). All tables can be found in the Appendix.

Different assumed annual income levels

The results for the relationship between income tax rates and residential location choice of inventors when I use the income tax rates for assumed annual income levels of 200,000 CHF, 300,000 CHF and 400,000 CHF can be found in Table 48 columns 5-10. The estimates of the elasticity of the probability to locate in a municipality with respect to the net-of-tax rate are 2.76 for an annual income of 200,000 CHF, 4.65 for an annual income of 300,000 CHF

[64] I use the same instrumented tax rate as in the main specification, i.e. the average of the top income tax rates of all other municipalities in the same state that lie farther than 5 kilometers away.

and 4.88 for an annual income levels of 400,000 CHF. In columns 1-4, I show the results for specifications that use the income tax rates that apply to incomes of 100,000 CHF and 150,000 CHF annually. The estimates of the corresponding elasticities are only 1.87 and 1.91 respectively, and the tax rates appear to explain less of the variation observed in location choices. Hence, income tax rates in the highest income tax brackets appear more predictive of the inventor's choice of residence.

Excluding small states, the state of Zürich or mountainous regions

In Table 49 columns 1 and 2, I present the results when I exclude the states of Zug, Schaffhausen, Nidwalden, Obwalden, Appenzell Innerrhoden, Appenzell Ausserhoden and Uri, which were the states with the fewest municipalities in 2010, and all municipalities in other states that lie close to the state borders of these states. A concern is that in smaller states, the local political preferences in specific state border municipalities may have a larger influence on state tax policies. The estimate of the elasticity of the probability to locate in a municipality with respect to the net-of-tax rate is 4.53. In columns 3 and 4, I exclude the state of Zürich, which is the state with the highest number of inventors, and all municipalities that lie close to the state borders of Zürich. The estimate of the elasticity of the probability to locate in a municipality with respect to the net-of-tax rate is 3.64 in this sample. When I exclude municipalities in mountainous regions, which are 10 percent or more covered by rocks, the estimate of the elasticity of the probability to locate in a municipality with respect to the net-of-tax rate is revised to 3.83 (columns 5 and 6).

Alternative definitions of inventors' choice sets

In Table 50 columns 1 and 2, I show the results when I expand the choice set of the inventor to municipalities within 7.5 km of distance to the inventor's municipality of residence. I define a municipality as a "state border" municipality if it lies within 7.5 km of distance of another municipality that belongs to a different state. Apart from enlarging the choice sets, this also enlarges the set of "state border municipalities" SB, defined in section 2.4, and consequently enlarges the set of inventors that are included in the analysis. In this specification, the estimate of the elasticity with respect to the net-of-tax rate is 3.67. An alternative way of defining the choice set of the inventor is to let the choice set be given by the municipalities that border the municipality of residence, instead of using the distance between the centroids of the municipalities to define "close-by" municipalities.[65] This leads, again, to slightly different choice sets and a different set of "state border municipalities". The results for this robustness check, which can be found in Table 50 columns 5 and 6, are very similar to the main results.

Inventors in urban centers

As explained in section 2.4.2, an important consideration is whether the results presented apply only if the inventor chooses to reside close by a state border. If preferences are heterogeneous, then inventors living far from a state border may exhibit preferences that are systematically different from inventors living close by a state border. This consideration is especially relevant for inventors who live in urban centers, which do not lie at state borders. In order to include as many inventors as possible (while keeping the computational burden manageable), I do the following: first, I extend the maximal distance between the inventor's municipality of residence and the other municipalities included in the choice sets to 20 km. Then, if there is a municipality within 20 km of distance of the inventor's municipality of residence that lies in a different state, I include the inventor in the sample. The sample comprises 96 percent of all inventors, in particular all inventors who reside in the urban centers

[65] One particular concern in this regard is that a municipality that is actually located at the border of a state but has a very large surface area, may be dropped from the set of "state border municipalities" SB_{5km} simply because there may be no other municipalities within 5 kilometers of distance of the centroid of the municipality

Zürich, Basel, Lausanne and Geneva.[66] The results are presented in Table 50 columns 3 and 4. The estimate of the elasticity of the probability to locate in a municipality with respect to the net-of-tax rate is 4.16 in the model that includes controls.

It is noteworthy that, in this robustness check, the relationship between top income tax rates and the residential location choice of inventors disappears in the parsimonious model without control variables (column 3). This findings suggests that the heterogeneity between municipalities swamps the effect of income taxes when comparing municipalities across different states that are up to 20 kilometer apart. In contrast, I find a strong relationship for municipalities that lie no farther than 5 kilometers apart in the parsimonious model of the main specification (Table 11 column 1).

2.5.2 Supporting evidence from changes in top income tax rates

In this section, I present evidence that corroborates the findings on the relationship between top income tax rates and residential location choices of inventors from the cross-section presented above. Since coverage of local income tax rates at the municipality-level between 2005 and 2009 is low in my dataset, I exploit changes in state tax policies. Beginning in 2008, the states of Thurgau and Sankt Gallen implemented state-wide reductions in local income taxes by reducing the rates of the state-component. I examine the effect of these state-wide tax cuts, which lead to an average decline in local top income tax rates of 3.1 and 3.3 percentage points respectively between 2007 and 2012 (in municipalities for which data is available), on the change in relative choice probabilities between municipalities on different sides of state borders over time. When an inventor appears on either side of a state border for the first time, I record in which state the inventor resides.[67] If the tax-cut indeed raised

[66] I only exclude inventors who reside in municipalities which do not lie within 20 km of distance to a municipality of a different state.

[67] The residential location of an inventor is observed whenever the inventor appears on a patent application.

the attractiveness of Thurgau and Sankt Gallen relative to its neighboring states, I expect that inventors are relatively more likely to appear in border municipalities that belong to Thurgau or Sankt Gallen in the later years of the sample. Using only state border municipalities, as opposed to using municipalities from the entire state, has the advantage that the urban city center of Zürich, which lies in the center of the neighboring state of Zürich, is not required to be on the same time trend as municipalities in the more rural states of Thurgau and Sankt Gallen. The urban city center of Zürich may have become more dominant in innovation in the larger region during this period irrespective of changes in income tax rates, for example, due to agglomeration effects.

I estimate a panel model of top income tax rates and the residential location choice of inventors, while maintaining my focus on state border municipalities. I use the subset of state border municipalities, which are defined in section 2.4, around the state borders of Thurgau and Sankt Gallen. Because there are not enough inventors to reliably estimate municipality-fixed effects, I aggregate municipalities to "state border regions". A state border region consists of all municipalities that belong to the same state and lie around the same state border. Each border region is paired with a corresponding border region that is part of the neighboring state. For example, the state border municipalities in Sankt Gallen at the border to Zürich constitute one border region, while the municipalities on the Zürich-side of the same border constitute another. I assign an inventors to a state border region when the inventors appears on either side of the state border for the first time. The location of the inventor is observed whenever the inventor is listed on a patent application. I do not track subsequent appearances of the same inventor (from subsequent patents) in the same state border region or in the corresponding border region on the other side of the state border in order not to confound my analysis with issues of inventor prolificness. In total, I consider 18 different state border regions. More details can be found in the Appendix.

As mentioned above, local income tax rates are missing for a substantial number of municipalities between 2005 and 2009. For each state that borders Sankt Gallen or Thurgau, I construct a panel of municipalities for which local income tax rates are available for all years between 2005 and 2012. I then use the average change in this subset of municipalities to proxy for state-wide changes in local income taxes. In the Appendix, I provide more details on the implementation.

I assume that the utility of inventor i from choosing state border region s in year t is given by the multinomial logit model

$$u_{ist} = \beta \log(1 - \bar{\tau}_{st}) + \gamma_s + \epsilon_{ist} \tag{14}$$

where $\bar{\tau}_{st}$ is the average top income tax rate in year t across all municipalities for which tax rates are available (throughout the entire period of 2005 to 2012) in the state of border region s, γ_s is a state border region fixed effect and ϵ_{ist} is the extreme-value distributed error term. In analogy to the empirical strategy employed in the main section, the choice set of the inventor only consists of the state border region which the inventor actually resides in and the corresponding border region on the other side of the state border. The set of state border region fixed effects is meant to absorb the effect of time-invariant factors that impact the relative attractiveness of the different sides of a state border, like differences in average commuting time. More details on the estimation can be found in the Appendix.

The results from estimating model (6) are presented in Table 13. The point estimates suggest that, as local income tax rates declined in Sankt Gallen and Thurgau relative to the neighboring states in later years, inventors were relatively more likely to appear in state border municipalities in Thurgau or Sankt Gallen. The point estimate of the elasticity of the probability to locate in a state border region with respect to the net-of-tax rate is 6.73 (column 1). However, the estimate is statistically insignificant (p=0.24). In the specification shown in column 1, I only consider municipalities that lie within 5 kilometers of distance to a municipality that belongs to a different state. In column 2, I show the results when I expand the maximal distance to 7.5 kilometers, thereby increasing the sample of inventors from 788 to 1187. Using the expanded sample, the point estimate of the elasticity of the probability to locate in a state border region with respect to the net-of-tax rate is 6.94. The 95-percent confidence interval of the estimate shrinks when I use the expanded sample, but continues to include 0. The associated p-value is 0.16 in a two-sided test. Overall, the results in this section are consistent with the findings from the main section.

2.5.3 Results on other amenities

I use the results from the model presented in Table 14 column 1, unless otherwise stated, to compare the relative magnitudes of the estimated coefficients. The considered specification includes the top income tax rate linearly (as opposed to the net-of-tax rate) and is based on the actual top income tax rate in the municipality (as opposed to the instrumented tax rate[68]). Since all variables in equation (9) enter the utility of the inventor linearly when I choose a linear specification for the top income tax rate, I am free to express the size of any coefficient in terms of the top income tax rate.

Commuting distance appears to have a dominant effect on the choice of the residential location. The increase in the number of inventors associated with a shorter distance between a municipality and the workplace by 1 kilometer corresponds to the effect of a reduction in the top income tax rate by 2.42 percentage points. When I consider the restricted sample,

[68] An updated version will contain all results based on instrumented tax rates.

which only includes inventors who live no further than 80 km away from their employer, the estimated coefficient of commuting distance is revised upwards (in absolute size) and found to correspond to a 2.52 percentage points decrease in the top income tax rate (column 3). Furthermore, the estimates from column 4 suggest that the marginal effect of reducing the distance to the workplace by one kilometer is halved after 30 km of distance (compared to the marginal effect of saving the "initial" kilometer of commuting distance), and becomes negligible after 60 km of distance. However, it is important to keep in mind that this distance is measured as a synthetic "straight-line" distance that does not take into account actual roadways and therefore understates differences in driving distance. The estimate of the effect of commuting distance is sensitive to including inventors that live in urban centers, which is considered as a robustness check in the robustness section.

The effect of a municipality having a *lakefront* is comparable in magnitude to the effect of a 5.25 percentage point lower top income tax rate. The effect of a shorter *distance to the national border* by 1 kilometer is found to correspond in magnitude to the effect of having a 0.34 percentage points lower top income tax rate. The main reason why it is desirable to be close to the national border is that prices are considerably lower in German, French and Italian border towns.[69]

Descriptively, I find that inventors appear to avoid municipalities that have a low population count. An increase in the *municipality population size* by 1,000 residents raises the number of inventors in that municipality on average in the same extent as a 1.20 percentage point decrease in the top income tax rate (column 1). The dependence of inventor utility on municipality population size appears to be non-linear, as can be seen in column 2. The

[69] Personal anecdotes suggest that it is a common practice for residents of Zürich to buy groceries at the German border.

marginal increase in the number of inventors in a municipality due to larger population is highest for very small municipalities, halved at a municipality population size of 12,000 and disappears at a population size of 24,000. Since there is only one municipality with a population size of more than 25,000 in my sample, I cannot estimate the dependence of residential location choices on population size outside of this range.

Sorting by language and political attitudes

In this section, I present evidence on how inventors' location choices correlate with language and attitudes towards immigration in municipalities, while controlling for income tax rates and all other variables used in section 2.5.1. I maintain the same research design and methodology.

The linguistic fragmentation of Switzerland allows me to study whether the location choices of foreign-born inventors from German and French speaking countries (i.e. from Germany, Austria or France) are aligned with the primary languages spoken in municipalities. Table 15 presents results on how inventors sort based on the share of German and French speakers in municipalities. The point estimates suggest that foreign-born inventors are more likely to sort into municipalities with lower shares of German speakers, unless the foreign-born inventors themselves stem from German speaking countries. However, the estimated coefficients are not statistically significant. A lower share of German speakers in a municipality may indicate a more cosmopolitan, or more diverse, environment.

In Table 16, I present results on how the inventors' residential location choices correlate with the vote share of the anti-immigration party SVP ("Schweizerische Volkspartei") in the federal elections of 2011. The underlying hypothesis is that the SVP vote share elicits the prevalence of "intolerance" in municipalities, which may be relevant to both, domestic and foreign-born inventors. I have included examples of election posters by the SVP in the Appendix to convince readers of the controversial nature of the party. That creative, inventive types may place particular value on “tolerance” as a location amenity was put forward by Richard Florida in his highly influential book “The Rise of the Creative Class” (2002). I find that the SVP vote share does indeed predict location choices of foreign-born and domestic inventors and both groups are more likely to reside in municipalities with a lower SVP vote share. The estimated coefficient appears larger for domestic inventors, which may be explained by domestic inventors being better acquainted with the values prevalent in municipalities. To explore this issue further, I include additional controls for the socioeconomic composition of municipalities. When I include the social status index, which is a measure of the share of individuals who are either university-educated, employed in a management position or earn a particularly high income, the coefficient of the SVP vote share stays negative for the pooled sample of inventors, but is rendered marginally insignificant (p=0.11). For domestic inventors, the coefficient is negative and marginally significant (p=0.07).

2.6 Discussion

2.6.1 Summary

The evidence presented in this chapter documents a strong relationship between the residential location choice of inventors and local top income tax rates in Switzerland. I find an elasticity of the number of inventors in a municipality with respect to the net-of-tax rate of around 4.6. This estimate is considerably higher than the elasticities found by Akcigit, Baslandze and Stantcheva (2016) and Moretti and Wilson (2017), who study the effect of personal income tax rates on inventor mobility between eight developed countries and between US states respectively. Since tax differences between neighboring municipalities may be exploited by relocating over short distances, it seems plausible that local tax policies have particularly strong effects in my setting. The results are consistent with supporting evidence from changes in income tax rates. Foreign-born inventors exhibit an elasticity with respect to the net-of-tax rate that is 44 percent higher than the elasticity for domestic inventors, although this difference is not statistically significant.

In addition, I asses the importance of other amenities in relation to the top income tax rate. I study how the residential location choice of inventors depend on amenities such as shorter commute, lakefronts and language. Furthermore, I find, consistent with Richard Florida's hypothesis that creative types place particular value on "tolerance" (Florida 2002), that inventors appear to have a bias against municipalities with stronger anti-immigration attitudes.

2.6.2 Further discussion: relevance of the location of residence and dynamic relocation

In this paper, I have taken the workplace of the inventor as given and focused exclusively on the choice of residence. At first glance, it seems evident that in order for a municipality (or for a region more generally) to thrive, it should host high-tech firms and provide employment for inventors. The workplace is assumed to be the place where knowledge is exchanged and where complimentary investments are made. In contrast, the location of the residence of the inventor is typically less of a concern for policymakers.

However, a longer distance between the municipality of residence and the workplace may make it more likely that the inventor, over time, transfers to a different workplace. Suppose, for example, that the initial inventor-employer relationship is disrupted and that the inventor is looking for a new employer. If the inventor prefers having a shorter commute, the inventor gives preference to employers located in municipalities near by. Thus, if the inventor was initially located close to his or her workplace, the inventor is more likely to accept a job in the same municipality. Relatedly, the inventor may be more likely to start a firm around the place where he or she lives. The data lends support to this hypothesis. Scatterplots of the relationship between the commuting distance to the initial workplace and the probability of switching workplace are presented in Figure 25.[70] The estimates obtained from a simple OLS Linear Probability Model (Table 17) document an economically large and statistically significant association. For example, an inventor who resides 25 kilometers away from his or her workplace is 49 percent more likely to switch workplace than an inventor for whom the municipality of residence and the workplace coincide (column 3).[71] For an inventor living 50 kilometers away, the corresponding increase in the probability to switch workplace is 129 percent (column 2). To corroborate the evidence, I test whether inventors transfer to a workplace that lies closer to home. In a one-sided test, I reject the hypothesis that inventors switch to a workplace that lies further away from their residence ($p = 0.03229$, not shown).

This mechanism of dynamic inventor relocation suggests that regions may benefit over time in their development if they are attractive as residential locations. A famous example is Zug, a rural state 30 kilometers south of Zürich, whose advent in innovation is often ascribed to its low-tax regime.[72] Due to its low income taxes and its proximity to the urban innovation center of Zürich, it is a popular location of residence for inventors employed in Zürich. Over

[70] A change of workplace is observed if the inventor appears at more than one workplace in the data set, i.e. if he or she appears as an inventor on patents filed by applicants that are situated in different municipalities. To be clear, if the inventor switches between two employers that are located in the same municipality, this does not constitute a change of workplace. Since I have not disambiguated the employers in my data set, it is unclear whether the inventors switch to a different establishment of the same firm, or change the employer altogether.

[71] Because not all changes of workplace are observed (as they are only recorded if patents are filed at both workplaces), the change in the probability to switch workplace associated with longer initial commuting distance is underestimated if measured in percentage points instead.

[72] See, for example, "Low tax Zug aims to become Switzerlands Crypto Valley", Reuters 10/08/2016 (linked), "How Switzerland became the Silicon Valley of Robotics", Forbes 10/26/2017 (linked) or "Welcome to Zug: the sleepy Swiss town that became a global hub", The Guardian 05/30/2008 (linked)

time, inventors transfer to workplaces in Zug, thereby increasing innovative activity in the region. The proximity to Zürich may therefore help explain how Zug managed to become an innovation hub, in contrast to other low-tax regions that have not seen similar increases in innovative activity. [73]

The findings in this chapter suggest that inventors are highly responsive to differences in top personal income tax rates at the local level. The results are potentially important for the design of tax policies in settings where individuals may be able to take advantage of tax differences by relocating over short distances. My findings may be particularly relevant in the context of the European Union: as legal and cultural barriers to the movement of inventors are receding, regions that border low-tax countries may find themselves at the risk of being stunted by "brain-drain". Many Western European countries, like the Benelux states, have densely populated border regions. The differences between the tax systems of European countries are large, especially for high-income earner. Further research to study the effect of different policies, amenities and cultural factors on inventor mobility in varying settings are an important avenue for future research.

[73] Another reason for why it may be relevant to study the preferences of inventors for residential locations is the advent of modern information technologies. Virtual collaboration software facilitates the disentanglement of workplace and residence, which may allow residential location preferences to play out unfettered in the future.

3 The international political economy of R&D subsidies and the consequences of trade integration

3.1 Introduction

Most developed countries regularly subsidize business research and development. Subsidies are provided directly, allocated at the discretion of government agencies, and indirectly through tax credits. Aside from raising the rate of innovation, government support for business R&D in part aims at promoting domestic firms in sectors with foreign competition. Over the past decades, multilateral institutions, such as the WTO and the European Commission, have successively integrated national economies into international markets. At the same time, policies that support business R&D continue to be set at a national level, potentially interfering with policies set by other nations and without much effort of harmonization.

This chapter presents a model to analyze the endogenous choice of public support for business R&D by two governments in a quality ladder model of endogenous growth a là Aghion and Howitt (1992). Although research subsidies generate cross-border externalities, governments are assumed to choose subsidy rates in an individually rational and non-cooperative fashion to maximize their welfare. I use this model to study government funding of R&D by domestic technology follower firms, that are "behind" in the technological race, and technology leader firms. I compare the case when product markets are separated between countries to the case when product markets are integrated and domestic and foreign firms compete on a common product market.

I find that in this model, product market integration leads to an increase in aggregate research spending and an increase in the share of business R&D funded by government.

Importantly, I find that product market integration changes the allocation of research subsidies; technology follower firms receive higher subsidies under integrated product markets. Furthermore, subsidies are more evenly distributed, i.e. in contrast to the case of separated product markets, where subsidies are directed at technology leader firms, there is less discrimination in research subsidy rates offered to different firms in the same country under integrated product markets. However, government welfare is found to be lower under integrated product markets, which indicates that the increase in subsidies to technology follower firms may be inefficient.

In the empirical section, I investigate whether increases in trade integration between 24 OECD countries between 1980 and 2006 are related to changes in government funding of business R&D and aggregate research spending. Consistent with the model predictions, the estimates suggest a positive, albeit statistically insignificant, relationship between trade openness and research spending, and a positive relationship between trade openness and the generosity of government support for business R&D. At the same time, there is a statistically significant relationship between trade openness and the relative generosity of indirect, indiscriminate R&D tax credits and direct R&D subsidies. My findings suggest that, as countries become more open to trade, they tend to switch from direct R&D subsidies to R&D tax credits. Since R&D tax credits tend to discriminate less between firms than direct subsidies, and its benefits are more evenly distributed and diffuse, this pattern appears consistent with the model.

The model presented in this chapter may in part explain a major shift in innovation policy globally over the last decades, namely the trend away from direct, discriminate R&D subsidies for business R&D towards indirect, indiscriminate R&D tax credits. This policy trend has received attention by the OECD, which tracks innovation policy in its member countries.[74] However, it is not quite clear why this shift has occured in OECD countries and why it might be related to changes in trade integration. Figure 13 shows the evolution of the share of business R&D funded directly by government for 18 OECD countries from 1980 to 2006. For the median country (in terms of share of business R&D funded directly by government) this share declined from around 12 percent to around 6 percent. At the same time, the implied tax subsidy rate by R&D tax credits has risen to around 8 percent for the median country, as evident in Figure 14. The relative generosity of direct government funding for business R&D and indirect R&D tax credit, measured by the difference in the implied subsidy rates per dollar spent on R&D, is shown in Figure 15. In the 18 OECD countries included in this panel, R&D tax credits have overtaken direct research subsidies in importance.[75]

---

[74] OECD Science and Technology Outlook 2014, p.169:

> "The general trend over the past decade has been to increase the [..] availability, generosity and simplicity of use of R&D tax incentives in the OECD area. Countries have redesigned their tax arrangements to make them more generous and attractive by raising thresholds on R&D expenditures and tax concessions or by increasing deduction rates and enlarging eligibility criteria. Many countries have abandoned incremental design for volume-based schemes that are simpler to implement for tax authorities and simpler to adopt for firms. As a consequence, public funding allocated to business R&D through tax incentives has increased markedly and R&D tax incentives have become a major instrument of STI policy in many countries."

The OECD started, beginning with the Science and Technology Outlook 2012, to include assessments of the relative importance of different types of research subsidies. The most important distinction is made between "direct government funding", a spending category that includes discretionary instruments like grants and procurement, and R&D tax credits, also referred to as "indirect government funding". The OECD describes the prime difference as follows:

> "Direct funding allows government to target specific R&D activities and steer business efforts towards new R&D areas that offer high social returns but low prospects for profit[..] Direct funding instruments depend on discretionary decisions by governments. Tax incentives reduce the marginal cost of R&D and innovation spending; the are usually more neutral than direct support in terms of industry, region and firm characteristics, although this does not exclude some differential, most often by firm size."(OECD Science and Technology Outlook 2014, p.156)

[75] The OECD has published results of similar calculations for the years 2012-2014, comparing estimates of forgone tax revenues to total direct government funding. Unfortunately, we are not able to replicate such calculations for our sample.

This chapter is organized as follows: Section 3.1.1 discusses related literature, section 3.2 presents the model and section 3.3 describes stylized facts about trade integration and government support for business R&D. Finally, section 3.4 concludes.

3.1.1 Related Literature

The present chapter relates to a small literature on the international political economy of R&D subsidies. Impullitti (2010) analyzes how foreign competition affects the optimal choice of R&D subsidies in the US, using a quality-ladder model similar to model studied in this paper. Garcia Pires (2015) studies a model in which competing technology follower firms and technology leader firms are hosted by different countries, and receive subsidies from their respective governments. Kondo (2012) studies how countries facilitate or prevent relocation of R&D intensive firms by offering R&D subsidies. R&D policy competition between countries that trade is also studied in Haaland and Kind (2008).

Finally, Hammadoua, Paty and Savonad (2014) study 14 European countries from 1996 and 2006 and find that countries that experienced stronger increases in trade openness increased their public R&D spending by more.

3.2 The Model

In this section, I present a framework to analyze the choice of research subsidy rates offered to domestic firms by two governments. The firms located in these countries either compete on separated or integrated product markets. Although research subsidies generate cross-border externalities, governments are assumed to choose subsidy rates in an individually rational and non-cooperative fashion to maximize their welfare.

In the first scenario, presented in section 3.2.1, two governments host domestic firms that engage in research and compete for a domestic product market. Although product markets are separated between the two countries, there are technological spillovers, and firms learn from the advances made by firms in the other country. In the second scenario, presented in section 3.2.2, the firms hosted by the two governments compete for an integrated product market with other firms from both countries. In section 3.2.3, I compare the equilibrium research subsidy rates chosen by governments across the two scenarios. First, I show that in this model, if subsidies are set exogenously, outcomes are identical under separated and integrated product markets. Then, I show that if subsides are chosen by governments in equilibrium, outcomes differ markedly across the two scenarios. Most importantly, technology follower firms that aspire to become technology leaders are subsidized more heavily if product markets are integrated.

3.2.1 Separated product markets

There are two countries A and B that each host two domestic firms. All firms operate in the same technological sector, but domestic firms do not compete with foreign firms in the product market. Firms undertake research to improve the quality of the only good that the sector produces in the vein of Aghion and Howitt (1991). Due to technological cross-border spillovers, there is a global technological state-of-the-art quality q for both countries.

In each country there is exactly one domestic technology leader and one technology follower firm at any point in time. The firm that is the current technology leader serves the entire domestic product market and earns a flow profit of $\beta q = \beta_A q = \beta_B q$. β represents the size of the domestic market.[76] The firm that is the current technology follower earns no flow

[76] The parameter β for the size of the market is a standard parameter in endogenous growth models with iso-elastic preferences or CRS final good production.

profit. The discounted value (payoff) of firm i in country $j \in \{A, B\}$ is given by

$$v_i^j(q_0) = E[\int_0^\infty exp(-rt)(\beta 1_{\{i \text{ is technology leader in j}\}} - (1 - s_{it})z_{it})q_t dt] \tag{15}$$

where r is the exogenous global interest rate, z_{it} is the research effort of the firm (which has a unit cost), s_{it} is the share of the cost of research covered by government funding (the "subsidy rate") and q_t is the quality state at time t. I denote by $v_L^j(q_0)$ the value of the current technology leader and by $v_F^j(q_0)$ the value of the current technology follower in country j. Note that in the definition of value (1), a firms only collects flow profits when it is in the technology leader position but may pay for exerted research effort when it is either in the leader or in the follower position. The firms understands that, even if it is the technology leader at the moment, it may in the future also spend some time in the technology follower position. All firms in both countries run independent but equally productive research lines that yield a successful innovation at the (stochastic poisson arrival) rate $\gamma(z) = 2\sqrt{z}$. If any firm in either country innovates, the global quality state is raised to λq, where $\lambda > 1$. Furthermore, I assume that if a firm innovates that is in the technology follower position, it becomes the new technology leader domestically. I assume that innovation by foreign firms does not lead to a switch in domestic leader and follower position.

Governments receive a flow payoff that is composed of three parts: the sum of the values of all domestic firms, the cost of research subsidies that it pays to domestic firms and an extra benefit that scales at the same rate by which quality and profits scale. This extra benefit to the government is not internalized by the firm. Examples could be corporate income tax revenue, wages paid to employees or consumer surplus. The discounted payoff of the government of country $j \in \{A, B\}$ is given by

$$g^j(q_0) = E[\int_0^\infty exp(-rt)(\beta + w - \sum_{i \in D_j} z_{it})q_t dt] \tag{16}$$

where D_j is the set of domestic firms in country j, w is the extra benefit to the government of country j. Note that

$$g^j(q_0) = \sum_{i \in D_j} v_i^j(q_o) + E[\int_0^\infty exp(-rt)(w - \sum_{i \in D_j} s_{it}^j z_{it})q_t dt] \tag{17}$$

The research subsidy rate s_{it}^j offered to domestic firm i by the government of country j is defined as follows: first, the government selects $s_{it}^j \geq 0$. Then, the firm selects a research effort z_{it}. If we assume a unit cost for research, this also constitutes the total cost of research. Of the total cost of research, the government covers $s_{it}^j z_{it}$ and the firm bears the remaining cost $(1 - s_{it}^j)z_{it}$.

In this model, governments may pay research subsidies because firms do not consider the benefit of their research efforts to the other domestic firm and firms do not consider the extra benefit of quality improvements to the government. However, this does not necessarily imply that the government chooses to subsidize all firms; technology follower firms may in fact be over-incentivized from a social perspective by the prospect of stealing business from the technology leader firm. Importantly, governments themselves do not consider the effect that their research subsidies have on the firms or the government in the other country.

Equilibrium

I assume that all firms and all governments select their research efforts and research subsidy rates non-cooperatively. I focus on symmetric markov perfect equilibria. Firms select their research efforts based on the aggregate state of quality and on the domestic competitive position they are in. Governments select research subsidy rates for domestic firms based on the aggregate state of quality. A markov perfect equilibrium is a tuple of value functions for firms $(v_L^j(\cdot), v_F^j(\cdot))_{j\in\{A,B\}}$, government value functions $(g^j(\cdot))_{j\in\{A,B\}}$, research effort by firms $(z_L^j(\cdot), z_F^j(\cdot))_{j\in\{A,B\}}$ and research subsidy rates $(s_L^j(\cdot), s_F^j(\cdot))_{j\in\{A,B\}}$. The indices L and F refer to the firms currently in the technology leader and in the technology follower positions. To make the notation more compact, I denote the research intensity of firm i in country j by

$$\gamma_i^j(q) = \gamma\big(z_i^j(q)\big) \tag{18}$$

and the change in value of firm i in country j if the global quality state is increased by one step (through innovation) when there is no switch in domestic technology leader and technology follower positions by

$$\Delta v_i^j(q) = v_i^j(\lambda q) - v_i^j(q) \tag{19}$$

The change in the value of the government of country j whenever there is innovation is denoted by

$$\Delta g^j(q) = g^j(\lambda q) - g^j(q) \tag{20}$$

In equilibrium, the following set of Hamilton-Jacobi-Bellman-equations must be satisfied in both countries $j \in \{A, B\}$:

$$\begin{aligned} rv_L^j(q) = \beta q + max_{z_L^j(q)\geq 0}\Big\{\gamma_L^j(q)\Delta v_L^j(q) - (1-s_L^j(q))z_L^j(q)q\Big\} - \gamma_F^j(q)\big(v_L^j(q) - v_F^j(\lambda q)\big) \\ + \big(\gamma_L^{-j}(q) + \gamma_F^{-j}(q)\big)\Delta v_L^j(q) \end{aligned} \tag{21}$$

$$rv_F^j(q) = max_{z_F^j(q)\geq 0}\Big\{\gamma_F^j(q)\big(v_L^j(\lambda q) - v_F^j(q)\big) - (1 - s_F^j(q))z_F^j(q)q\Big\} + \gamma_L^j(q)\Delta v_i^j(q) + \big(\gamma_L^{-j}(q) + \gamma_F^{-j}(q)\big)\Delta v_F^j(q) \tag{22}$$

$$\begin{aligned} rg^j(q) = (\beta + w)q &+ max_{s_L^j(q)\geq 0}\Big\{\gamma_L^j(q)\Delta g^j(q) - z_L^j(q)q\Big\} \\ &+ max_{s_F^j(q)\geq 0}\Big\{\gamma_F^j(q)\Delta g^j(q) - z_F^j(q)q\Big\} \\ &+ \big(\gamma_L^{-j}(q) + \gamma_F^{-j}(q)\big)\Delta g^j(q) \end{aligned} \tag{23}$$

Note that the research intensity of firm $\gamma_i^j(q)$ is a function of the research effort $z_i^j(q)$, as evident in definition (4).

Finding functions that satisfy equations (21) – (23) is straightforward: the flow payoffs of all firms and all governments are linear in the quality state, which grows at a constant step size λ. This implies that in a markov perfect equilibrium, all value functions are linear in the quality state q. Hence, there exist scalars s.t. $g^j(q) = g^j q$, $v_L^j(q) = v_L^j q$ and $v_L^j(q) = v_L^j q$ for both countries $j \in \{A, B\}$. Furthermore, in a symmetric equilibrium the value functions are identical for both countries. It is then easy to verify that the equilibrium research efforts and research subsidy rates (z_L, z_F, s_L, s_F) do not depend on the quality state q. Definitions (4) – (6) become

$$\gamma_L = \gamma\big(z_L\big),\ \gamma_F = \gamma\big(z_F\big) \tag{24}$$

$$\Delta v_L = v_L\lambda - v_L,\ \Delta v_F = v_F\lambda - v_F \tag{25}$$

and

$$\Delta g = g\lambda - g \tag{26}$$

Equations (7) – (9) become

$$rv_L = \beta + \big(\gamma_L\Delta v_L - (1 - s_L)z_L\big) - \gamma_F\big(v_L - v_F\lambda\big) + \big(\gamma_L + \gamma_F\big)\Delta v_L \tag{27}$$

$$rv_F = \big(\gamma_F(v_L\lambda - v_F) - (1-s_F)z_F\big) + \gamma_L \Delta v_F + \big(\gamma_L + \gamma_F\big)\Delta v_F \tag{28}$$

$$\begin{aligned} rg = (\beta + w) + \Big(\gamma_L \Delta g - z_L\Big) + \Big(\gamma_F \Delta g - z_F\Big) \\ + (\gamma_L + \gamma_F)\Delta g \end{aligned} \tag{29}$$

The conditions that ensure that firm research efforts and research subsidy rates are chosen optimally in equilibrium, assuming that the research lines of all firms are independent and yield innovation at the rate $\gamma(z) = 2\sqrt{z}$, are given by

$$z_L = \Big(\frac{\Delta v_L}{1-s_L}\Big)^2 \tag{30}$$

$$z_F = \Big(\frac{v_L\lambda - v_F}{1-s_F}\Big)^2 \tag{31}$$

and

$$s_L = max\Big\{1 - \frac{\Delta v_L}{\Delta g}, 0\Big\} \tag{32}$$

$$s_F = max\Big\{1 - \frac{v_L\lambda - v_F}{\Delta g}, 0\Big\} \tag{33}$$

I summarize below:

Definition 1. *A symmetric markov equilibrium of the game with separated product markets is a tuple of values* $(v_L, v_F, g) \in \mathbb{R}^3$ *and a tuple of research efforts, research subsidiy rates and research intensities* $(z_L, z_F, s_L, s_F, \gamma_L, \gamma_F) \in \mathbb{R}^6$ *that solve equations (24)-(33).*

Definition 2. *Consider a symmetric markov equilibrium of the game with separated product markets defined in Definion 1. Then,*

- *"Government welfare" in a symmetric equilibrium with separated product markets is defined as the payoff of the government* g

- *"Aggregate R&D spending" in a symmetric equilibrium is defined as*

$$z_L + z_F$$

- *"Share of aggregate R&D spending funded by government" in a symmetric equilibrium is defined as*

$$\frac{s_L z_L + s_F z_F}{z_L + z_F}$$

3.2.2 Integrated Product Markets

In this section, I assume that the firms compete on an integrated product market. To isolate the effect that the change in governments' funding for research have on firms and aggregate welfare, the integrated product market is modeled in a way that yields identical results in the absence of government funding for research. I assume that in the integrated product market, all firms from both countries compete on a common product market. However, this product market is segmented, and at any point in time there are exactly two technology leader firms and two technology follower firms. Each technology leader firm serves exactly one half of the entire product market. The setup closely resembles the case of separated product markets: each technology leader earns a flow profit of β and the value of firm i in country $j \in \{A, B\}$ is, exactly as in the case of separated product markets, given by

$$\tilde{v}_i^j(q_0) = E[\int_0^\infty exp(-rt)(\beta 1_{\{i \text{ is a technology leader}\}} - (1 - s_{it})z_{it})q_t dt] \tag{34}$$

where all definitions are is in (1). All firms run independent and equally productive research lines and if any of the firms innovates, the global state of quality q in the sector is raised to λq. Similar to the case of separated product markets, if a firm innovates that is in the technology follower position, it replaces a technology leader. In this case, one of the incumbent technology leader is chosen at random to remain in a leader position while the other technology leader becomes a technology follower.

The discounted payoff of the government of country j is given by

$$\tilde{g}^j(q_0) = \sum_{i \in D_j} \tilde{v}_i^j(q_o) + E[\int_0^\infty exp(-rt)(\tilde{w}_t^j - \sum_{i \in D_j} z_{it})q_t dt] \tag{35}$$

where, as before, D_j is the set of domestic firms in country j. I assume that governments only fund the research of domestic firms.[77] There is one difference between (21) and (2): when firms compete on an integrated product market, the extra benefit of the government of country j, $\tilde{w}_t^j$, depends on whether one or all of the technology leader firms are actually domestic firms. I will assume assume that $\tilde{w}_t^j = 0$ whenever country j hosts none of the incumbent technology leaders, $\tilde{w}_t^j = w$ if it hosts exactly one technology leader and $\tilde{w}_t^j = 2w$ if it hosts all.

Equilibrium

The discussion in this section is an extension of section 3.2.1. When product markets are integrated, the state space for firms and governments is enlarged. While there is, as before, only a single global state of quality q, governments may select research subsidy rates for domestic firms dependent on how many domestic firms are currently technology leaders. Consequently, firms' research efforts depend on the competitive position of the other domestic firm. From the perspective of any firm, the state space is $\{LL, LF, FL, FF\}$, where, for example, LF refers to the state where the firm is currently a technology leader, while the other domestic firm is a technology follower. From the perspective of the government, the

[77]If we assume that governments have at least an arbitrarily small bias in favor of domestic firms, this assumption is without loss of generality, in the sense that, in equilibrium, governments would only subsidze domestic firms.

state space is $\{LL, LF, FF\}$, where LL refers to the state where both domestic firms are technology leaders, LF refers to the state where exactly one of the firms is a technology leader and FF refers to the state where both domestic firms are technology followers.

A markov perfect equilibrium is then a tuple of value functions for firms $(v^j_{LL}(\cdot), v^j_{LF}(\cdot), v^j_{FL}(\cdot), v^j_{FF}(\cdot))_{j\in\{A,B\}}$, government value functions $(g^j_{LL}(\cdot), g^j_{LF}(\cdot), g^j_{FF}(\cdot))_{j\in\{A,B\}}$, research effort by firms $(z^j_{LL}(\cdot), z^j_{LF}(\cdot), z^j_{FL}(\cdot), z^j_{FF}(\cdot))_{j\in\{A,B\}}$ and research subsidy rates $(s^j_{LL}(\cdot), s^j_{LF}(\cdot), s^j_{FL}(\cdot), s^j_{FF}(\cdot))_{j\in\{A,B\}}$. As in the previous section, I restrict attention to symmetric markov perfect equilibria. The problem then reduces, as in the previous section, to finding scalars $(v_{LL}, v_{LF}, v_{FL}, v_{FF}, g_{LL}, g_{LF}, g_{FF})$ such that for all value functions, it holds true that $v^j_{LL}(q) = v_{LL}q$ and so on, and scalars $(z_{LL}, z_{LF}, z_{FL}, z_{FF}, s_{LL}, s_{LF}, s_{FL}, s_{FF})$ such that $z^j_{LL}(q) = z_{LL}$ and so on. To make the notation more compact, I define the research intensities and the change in values when there is no switch between technology leader and technology follower positions as $\Delta v_s = v_s\lambda - v_s$, $\Delta g_s = g_s\lambda - g_s$ and $\gamma_s = \gamma\big(z_s\big)$ respectively for all states s. In equilibrium, the following set of HJB-equations has to hold:

$$rv_{LL} = \beta + \big(\gamma_{LL}\Delta v_{LL} - (1 - s_{LL})z_{LL}\big) + \gamma_{LL}\Delta v_{LL} + \gamma_{FF}\big((v_{LF}\lambda - v_{LL}) + (v_{FL}\lambda - v_{LL})\big) \tag{36}$$

$$\begin{aligned} rv_{LF} = \beta + &\big(\gamma_{LF}\Delta v_{LF} - (1 - s_{LF})z_{LF}\big) + \gamma_{FL}\big(\frac{1}{2}(v_{FL}\lambda - v_{LF}) + \frac{1}{2}(v_{LL}\lambda - v_{LF})\big) \\ &+ \gamma_{FL}\big(\frac{1}{2}(v_{FF}\lambda - v_{LF}) + \frac{1}{2}\Delta v_{LF}\big) + \gamma_{LF}\Delta v_{LF} \end{aligned} \tag{37}$$

$$\begin{aligned} rv_{FL} = &\gamma_{FL}\big(\frac{1}{2}(v_{LF}\lambda - v_{FL}) + \frac{1}{2}(v_{LL}\lambda - v_{FL})\big) - (1 - s_{FL})z_{FL} + 2\gamma_{LF}\Delta v_{FL} \\ &+ \gamma_{FL}\big(\frac{1}{2}(v_{FF}\lambda - v_{FL}) + \frac{1}{2}\Delta v_{FL}\big) \end{aligned} \tag{38}$$

$$rv_{FF} = \big(\gamma_{FF}(v_{LF}\lambda - v_{FF}) - (1 - s_{FF})z_{FF}\big) + \gamma_{FF}(v_{FL}\lambda - v_{FF}) + 2\gamma_{LL}\Delta v_{FF} \tag{39}$$

$$rg_{LL} = 2(\beta + w) + 2\big(\gamma_{LL}\Delta g_{LL} - z_{LL}\big) + 2\big(\gamma_{FF}(g_{LF}\lambda - g_{LL})\big) \tag{40}$$

$$\begin{aligned} rg_{LF} = (\beta + w) &+ \big(\gamma_{LF}\Delta g_{LF} - z_{LF}\big) \\ &+ \big(\gamma_{FL}(\frac{1}{2}\Delta g_{LF} + \frac{1}{2}(g_{LL}\lambda - g_{LF})) - z_{FL}\big) \\ &+ \gamma_{LF}\Delta g_{LF} + \gamma_{FL}\big(\frac{1}{2}\Delta g_{LF} + \frac{1}{2}(g_{FF}\lambda - g_{LF})\big) \end{aligned} \tag{41}$$

$$rg_{FF} = 2\big(\gamma_{FF}(g_{LF}\lambda - g_{FF}) - z_{FF}\big) + 2\gamma_{LL}\Delta g_{FF} \tag{42}$$

The conditions that ensure that firm research efforts and research subsidy rates are chosen optimally in equilibrium, assuming that the research lines of all firms are independent and yield innovation at the rate $\gamma(z) = 2\sqrt{z}$, are given by

$$z_{LL} = \Big(\frac{\Delta v_{LL}}{1 - s_{LL}}\Big)^2,\ z_{LF} = \Big(\frac{\Delta v_{LF}}{1 - s_{LF}}\Big)^2 \tag{43}$$

$$z_{FL} = \Big(\frac{\frac{1}{2}\lambda(v_{LF} + v_{LL}) - v_{FL}}{1 - s_{FL}}\Big)^2,\ z_{FF} = \Big(\frac{v_{LF}\lambda - v_{FF}}{1 - s_{FF}}\Big)^2 \tag{44}$$

and

$$s_{LL} = max\Big\{1 - \frac{\Delta v_{LL}}{\Delta g_{LL}}, 0\Big\},\ s_{LF} = max\Big\{1 - \frac{\Delta v_{LF}}{\Delta g_{LF}}, 0\Big\} \tag{45}$$

$$s_{FL} = max\Big\{1 - \frac{\frac{1}{2}\lambda(v_{LF} + v_{LL}) - v_{FL}}{\frac{1}{2}\Delta g_{LF} + \frac{1}{2}(g_{LL}\lambda - g_{LF})}, 0\Big\},\ s_{FF} = max\Big\{1 - \frac{v_{LF}\lambda - v_{FF}}{g_{LF}\lambda - g_{FF}}, 0\Big\} \tag{46}$$

As before, the equilibrium of the game is defined as the solution to this set of equations.

Definition 3. *A symmetric markov equilibrium of the game with integrated product markets is a tuple of values* $(v_{LL}, v_{LF}, v_{FL}, v_{FF}, g_{LL}, g_{LF}, g_{FF}) \in \mathbb{R}^7$ *and a tuple of research efforts, research subsidiy rates and research intensities* $(z_{LL}, z_{LF}, z_{FL}, z_{FF}, s_{LL}, s_{LF}, s_{FL}, s_{FF}, \gamma_{LL}, \gamma_{LF}, \gamma_{FL}, \gamma_{FF}) \in \mathbb{R}^{12}$ *that solve equations (36)-(46).*

Under integrated product markets, the payoff of the government depends on the current state of competition. I will refer to the state where both countries host exactly one technology leader as the "leveled state" and to the state in which one country hosts both technology leaders as the "unleveled state". In the equilibrium under integrated product markets, the stationary distribution over leveled and unleveled states is given by

$$p_u = P(\text{"unleveled state"}) = \frac{1}{2 + \frac{\gamma_{FF}}{\gamma_{FL}}} \tag{47}$$

$$p_l = P(\text{"leveled state"}) = 1 - \frac{1}{2 + \frac{\gamma_{FF}}{\gamma_{FL}}} \tag{48}$$

I then define government welfare, aggregate R&D spending and R&D subsidies as a share of R&D spending analogous to the case of separated product markets:

Definition 4. *Consider a symmetric markov equilibrium of the game with integrated product markets defined in Definion 3. Then,*

- *"Government welfare" in a symmetric equilibrium with integrated product markets is defined as*

$$p_l g_{LF} + p_u\big(\frac{1}{2}g_{LL} + \frac{1}{2}g_{FF}\big)$$

- *"Aggregate R&D spending " in a symmetric equilibrium with integrated product markets is defined as*

$$p_u\big(z_{LF} + z_{FL}\big) + p_l\big(z_{LL} + z_{FF}\big)$$

- *"Share of aggregate R&D spending funded by government" in a symmetric equilibrium with integrated product markets is defined as*

$$\frac{p_l\big(s_{LF}z_{LF} + s_{FL}z_{FL}\big) + p_u\big(s_{LL}z_{LL} + s_{FF}z_{FF}\big)}{p_u\big(z_{LF} + z_{FL}\big) + p_l\big(z_{LL} + z_{FF}\big)}$$

3.2.3 A comparison of research subsidy rates and research spending under separated and integrated product markets

In this section, I compare government welfare, R&D spending and R&D subsidies as a share of R&D spending under separated and integrated product markets. In section 3.2.3, I show that outcomes differ because governments choose different research subsidy rates in equilibrium across the two scenarios. As benchmark result, I establish that, if subsidy rates are fixed across the two scenarios, outcomes are identical.

In contrast, I show that, if the choice of research subsidy rates is endogenous, governments adjust their research subsidy rates and outcomes are generally different under separated and integrated product markets. I simulate the model for a wide range of parameters. Across the entire range of parameters considered, government welfare is *lower* under integrated product markets, while aggregate R&D spending and R&D subsidies as a share of aggregate R&D spending are *higher*. Furthermore, I show that there is *less* discrimination in subsidy rates offered to firms under integrated product markets and that, more specifically, not only domestic technology leader firms, but also domestic technology follower firms receive considerable subsidies under integrated product markets.

Exogenous research subsidies

In the definition of the symmetric markov equilibrium for the game with separated product markets (definition 1), governments choose their research subsidy rates in an individually rational and non-cooperative fashion. Suppose that instead, subsidy rates (s_F, s_L) were fixed exogenously, while firms still choose their research efforts optimally given the subsidy rates they are offered. In this case, I may still use equations (27) - (31) to compute the equilibrium of this augmented game. Analogously, given exogenously fixed subsidy rates $(s_{FF}, s_{FL}, s_{LF}, s_{LL})$, I may use equations (36) - (44) to compute the equilibrium of the augmented game under integrated product markets.

Definition 5. *Consider a tuple of subsidy rates* $(\bar{s}_F, \bar{s}_L) \in \mathbb{R}^2$ *for the game with separated product markets, and a tuple of subsidy rates* $(\bar{s}_{FF}, \bar{s}_{FL}, \bar{s}_{LF,LL}) \in \mathbb{R}^4$ *for the game with integrated product markets.*

A symmetric equilibrium of the augmented game with separated product markets and fixed subsidy rates is a tuple of values $(v_L, v_F, g) \in \mathbb{R}^3$ *and a tuple of research efforts and research intensities* $(z_L, z_F, \gamma_L, \gamma_F) \in \mathbb{R}^4$ *that solve equations (24)-(31), where the research subsidy rates are given by* $(\bar{s}_F, \bar{s}_L)$.

Furthermore, a symmetric equilibrium of the augmented game with integrated product markets and fixed subsidy rates is s a tuple of values $(v_{LL}, v_{LF}, v_{FL}, v_{FF}, g_{LL}, g_{LF}, g_{FF}) \in \mathbb{R}^7$ *and a tuple of research efforts and research intensities* $(z_{LL}, z_{LF}, z_{FL}, z_{FF}, \gamma_{LL}, \gamma_{LF}, \gamma_{FL}, \gamma_{FF}) \in \mathbb{R}^{12}$ *that solve equations (36)-(44), where research subsidy rates are given by* $(\bar{s}_{FF}, \bar{s}_{FL}, \bar{s}_{LF}, \bar{s}_{LL})$.

Proposition 1 establishes the benchmark result that outcomes are identical if subsidy rates are fixed exogenously and set constant across the two scenarios.

Proposition 4. *Consider a tuple of fixed subsidy rates* $(\bar{s}_F, \bar{s}_L) \in \mathbb{R}^2$ *for the augmented game with separated product markets, and a tuple of fixed subsidy rates* $(\bar{s}_{FF}, \bar{s}_{FL}, \bar{s}_{LF}, \bar{s}_{LL}) \in \mathbb{R}^4$ *for the augmented game with integrated product markets (see definition 5). Suppose that subsidy rates are identical for technology follower and technology leader firms across the two scenarios, i.e. they satisfy*

$$\bar{s}_F = \bar{s}_{FL} = \bar{s}_{FF}$$

and

$$\bar{s}_L = \bar{s}_{LF} = \bar{s}_{LL}$$

Then, government welfare, aggregate R&D spending and R&D subsides as a share of aggregate R&D spending are identical under separated and integrated product markets.

Endogenous research subsidies

In this section, I contrast the outcomes under separated and integrated product markets when research subsidy rates are chosen by governments endogenously. Since closed-form solutions for the dynamic model described in section 3.2.2. are not available, I calculate equilibria for the games with separated product markets and integrated product markets numerically. The parameters for my preferred calibration are chosen as follows: First, I assume that the interest rate r and the flow profit β equal 1. I assume an increase in quality in each innovative step of 10 percent, which corresponds to a choice of λ of 1.1. The last parameter to be chosen is the external benefit of innovation that accrues to the government w. Given our choices of (r, β, λ), I choose $w = 0.8$ (meaning that the external benefit to the government is 80 percent of the flow profit that accrues to the firm) so that R&D subsidies as a share of aggregate R&D spending are close to 10 percent in equilibrium when product markets are separated (see Table 4), which is close to the average figure for countries analyzed in the empirical section in 1980.

The results are presented in Tables 2 - 11 for a large grid of parameters of λ and w. I discuss the results for my preferred calibration, noting that qualitatively the results are similar for all parameter values considered.

Spending on research subsidies

Table 20 shows the difference in aggregate research spending between the two scenarios. Aggregate research spending under integrated product markets is as much as 44.15 percent higher than under separated product markets in my preferred calibration.

In my preferred calibration, R&D subsidies as a share of aggregate R&D spending are drastically higher under integrated product markets. When product markets are integrated, R&D subsidies account for 31.99 percent of aggregate R&D spending, compared to only 9.96 percent when product markets are separated.

Allocation of research subsidies

Tables 6 - 11 show the research subsidy rates that are offered in equilibrium to firms under separated product markets and under integrated product markets. Recall that subsidy rates in my model are defined as follows: First, the government announces a firm-specific subsidy rate s. The firm then chooses its research effort z, which has a unit cost. The government pays for a share of sz of the total cost and the firm pays for remaining $(1 - s)z$.

Subsidies are paid for two reasons: First, firms do not internalize the benefit to the other domestic firm when they choose their research effort and, second, firms do not internalize the external benefit that accrues to governments. However, since firms are also partially motivated by (socially wasteful) business stealing, it is not a priori clear whether governments will choose to pay subsidies to all firms. Governments in turn do not internalize the benefit of their research subsidies to the other government and the firms that it hosts.

Table 23 shows that when product markets are separated, technology follower firms receive no subsidies. This suggests that technology follower firms are in fact over-incentivized to innovate by the prospect of business stealing from the government's perspective. Business stealing occurs only between domestic firms when product markets are separated, and, in contrast to the firms, governments do not value changes in competitive positions per se. In contrast, Table 24 shows that technology leader firms receive high subsidies. In my preferred calibration, the subsidy rate offered to technology leader is 65.4 percent, which means that government pays for almost two-third of research spending by technology leader firms.

When product markets are integrated, technology follower firms receive research subsidies for a large set of parameters, as can be seen in Tables 25 and 27. I refer to the state of competition between the countries as "leveled" if both countries host exactly one technology leader and one technology follower firm. If one country hosts both technology leader firms, and the other country hosts both technology followers, I call the state "unleveled". In both states of competition between the countries, leveled and the unleveled, technology follower firms are offered higher subsidy rates than under separated product markets. This finding is consistent with the notion that governments have less incentive to depress business stealing by technology follower firms when product markets are integrated; as business is stolen fully or partially from foreign firms, governments find it in their interest to subsidize technology follower firms. In my preferred calibration, technology follower firms are offered research subsidy rates of 9.25 and 64.32 percent in leveled and unleveled states of competition between the countries respectively.

Technology leader firms are offered higher subsidy rates when product markets are integrated in the unleveled state and lower subsidy rates in the leveled state of competition compared to the case of separated product markets, as evident in Tables 26 and 28. In my preferred calibration, technology leader firms are offered research subsidy rates of 45.65 and

73.47 percent in leveled and unleveled states of competition respectively.

Importantly, when product markets are integrated, there is less discrimination in the research subsidy rates offered to different firms in the same country relative to the case of separated product markets. If all firms in the same country are either technology follower or technology leader firms, they are offered the same research subsidy rates by their governments. If the state of competition is leveled, and each country hosts exactly one technology follower and one technology leader, follower firms are offered higher research subsidy rates compared to the rates offered to followers under separated product markets. At the same time, technology leader firms are offered lower subsidy rates compared to the rates offered to technology leaders under separated product markets, implying an overall decline in the extent of discrimination between technology leader firms and technology follower firms. This can be seen by comparing the subsidy rates shown in Table 26 and Table 25 for integrated product markets, and the subsidy rates shown in Table 24 and Table 23 for separated product markets. The difference between the subsidy rates offered to technology leader firms and the subsidy rates offered to technology follower firms is smaller when product markets are integrated.

Government welfare

Table 19 shows that government welfare is lower under integrated product markets than under separated product markets across the entire range of parameters considered. In my preferred calibration, government welfare is 2.41 percent lower under integrated product markets. This is surprising insofar as aggregate research spending is higher under integrated product markets. In this model, product market integration induces governments to funnel research subsidies to technology follower firms in a way that reduces overall welfare in equilibrium.

3.3 Data

In this section, I utilize a dataset on direct research subsidies for business R&D and R&D tax credits across 24 OECD countries from 1981-2006 to relate changes in government funding for business R&D to increased product market integration.[78] As a measure of product market integration, I use trade openness, which is available for all countries in my dataset and collected in a consistent way across all countries. In section 3.3.1, I explain the construction of the dataset. Section 3.3.2 presents several stylized facts on the relationship between government funding of business R&D and trade openness.

3.3.1 Construction of the dataset

The data used in the analysis is based on three sources: data on the generosity of R&D tax credits in 26 OECD countries from 1980-2006 was collected and published by Thomson (2012), data on direct research subsidies for business R&D and total business research expenditures is taken from the OECD and data on trade openness stems from the World Bank.

As measure for the generosity of R&D tax credits, I use the B-index. The B-index is a widely used user-cost measure of R&D expenditures (Warda 2001). It measures, given the current corporate income tax rate and R&D tax credit scheme, how much pre-tax revenue must be earned this period to earn back one dollar of R&D expenditure. Thomson (2012) defines the B-index as

$$B = \frac{1 - \text{deduction rate} * \tau - \text{extra deduction}}{1 - \tau} \tag{49}$$

[78]Countries included are: Australia, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Japan, Korea, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Spain, Sweden, UK, US.

where τ is the corporate income tax rate. Put differently, the B-index is the effective cost of one dollar of R&D expenditure to the firm.[79] The tax subsidy rate is defined as

$$TSR = 1 - B \tag{50}$$

The generosity of direct funding is measured as the share of business R&D expenditures funded directly by government, which I denote by DSR (for direct subsidy rate). To motivate my measures of relative and total generosity of R&D tax credits and direct funding, I would like to give this share the interpretation of a subsidy offered to a representative firm. I then propose the sum

$$SR = DSR + TSR \tag{51}$$

(SR for subsidy rate) as a measure of the total generosity of government funding for business R&D and the difference

$$RSR = DSR - TSR \tag{52}$$

(RSR for relative subsidy rate) as a measure of the relative generosity of R&D tax credits and direct funding, where a decrease indicates relatively more subsidization through tax credits. I consider business expenditures for R&D as a share of GDP, often referred to as "BERD", as a measure of aggregate spending on R&D by businesses. As mentioned, I use trade openness, which is defined as (Imports+Exports)/GDP, as my measure of international product market integration.

[79]One important caveat is that it is assumed that the firm has sufficient tax liabilities to claim the full credit. While this is a serious concern in theory, it is less so in practice as all countries in our sample have very generous carry-forward regulations. A second caveat is that in some countries the B-index depends on firm size, with preferential treatment given to small firms. I refer to Thomson (2012) for a discussion of how to aggregate to a single B-index for a country-year pair.

The panel includes 24 out of 26 of the countries considered by Thomson (2012).[80] Thomson's panel on R&D tax credits covers 23 out of 26 countries completely from 1980-2006 and Hungary, Poland and Czech Republic from the early 1990s onwards. However, data on the share of business R&D expenditures funded directly by government from the OECD is missing for a substantial share of country-year pairs. For some countries, extended periods of 5 years or more are missing, while others, like Sweden and Norway, are reported on a biannual basis. This is a serious problem because, even if data is missing at random, up to 35 percent of all observations are missing when I compute first differences. I therefore impute the share of business R&D expenditures funded by government for years for which the values of the preceding and the following year are available by the arithmetic average of the two values.[81] Overall, the dataset contains 24 countries and 539 country-year pairs. Lists of all included country-year pairs can be found in the Appendix.

3.3.2 Empirical Analysis

In this section, I test whether increases in trade openness are related to changes in total business R&D expenditures and changes in the relative and total generosity of R&D tax credits and direct government funding for business R&D. More specifically, I estimate the first-differenced models

$$\Delta \text{BERD}_{it} = \alpha \Delta \text{Open}_{it} + \sum_{j=1981}^{2006} \beta_t 1_{t=j} + \epsilon_{it} \tag{53}$$

$$\Delta \text{SR}_{it} = \alpha \Delta \text{Open}_{it} + \sum_{j=1981}^{2006} \beta_t 1_{t=j} + \epsilon_{it} \tag{54}$$

[80] I drop Austria and Switzerland since business expenditures are only reported for every forth or fifth year.

[81] For robustness, I also consider a panel of 20 countries and 396 country-year pairs without imputations. Results do not depend on imputations.

and

$$\Delta \text{RSR}_{it} = \alpha \Delta \text{Open}_{it} + \sum_{j=1981}^{2006} \beta_t 1_{t=j} + \epsilon_{it} \tag{55}$$

where i is the country-index and t is the time index. The models include a full set of year dummies which control for common trends and year-by-year changes across countries. As explained in the previous section, $BERD$ measures business expenditures for R&D as a share of GDP, *Open* is defined as (Imports+Exports)/GDP, RSR is a measure of the relative generosity of R&D tax credits and direct funding (a decrease indicates relatively more subsidization through tax credits) and SR measures the total generosity of government funding for business R&D through tax credits and direct funding.

Empirical results

Estimation results for regression equations (53) - (55) are presented in table 18. The point estimate of the relationship between business research expenditures $BERD$ and trade openness, and the point estimate of the relationship between the total generosity of direct and indirect government funding for business R&D SR and trade openness are both positive, but statistically insignificant. The model analyzed in Section 3.2 predicts a positive relationship between aggregate research spending and product market integration, and a positive relationship between research subsidies as a share of research spending and product market integration.

The estimate of the relationship between the relative generosity of direct and indirect government funding for business R&D RSR and trade openness is negative and statistically significant. This finding suggests that, as countries become more open to trade, they tend to switch from direct funding of business R&D to indirect R&D tax credits. The model analyzed in Section 3.2 predicts that product market integration leads to an increase in research subsidies to technology follower firms and an overall decline in the extent of discrimination between firms in the same country. As R&D tax credits constitute a more or less indiscriminate way of subsidizing business R&D, that distributes subsidies more evenly across firms in the same country, the empirical findings and the model appear consistent.

My estimates suggest that an increase in trade openness of around 30 percentage points is associated with a shift in the relative generosity of R&D tax credits and direct research subsidies by around 4.5 cent per dollar towards R&D tax credits.

3.4 Conclusion

This chapter presents a framework to analyze the endogenous choice of research subsidies by two governments. Although research subsidies generate cross-border externalities, governments are assumed to choose subsidy rates in an individually rational and non-cooperative fashion to maximize their welfare. I use this model to compare research subsidies paid to domestic technology follower and technology leader firms when product markets are separated between countries to the case when product markets are integrated and domestic and foreign firms compete on a common product market.

I find that in this model, product market integration leads to an increase in aggregate research spending and an increase in research subsidies as a share of aggregate research spending. Importantly, I find that product market integration changes the allocation of research subsidies; technology follower firms receive higher subsidies under integrated product markets. Furthermore, subsidies are more evenly distributed, i.e. there is less discrimination in research subsidy rates offered to firms in the same country under integrated product markets. However, government welfare is found to be lower under integrated product markets, which indicates that the increase in subsidies to technology follower firms may be inefficient.

In the empirical section, I investigate whether increases in trade integration between 24 OECD countries between 1980 and 2006 are related to changes in research subsidies and research spending. Consistent with the model predictions, the estimates suggest a positive relationship between trade openness and research spending, and a positive relationship between trade openness and the generosity of government support for business R&D. However, these results are not statistically significant. On the other hand, there is a statistically significant relationship between trade openness and the relative generosity of indirect, indiscriminate R&D tax credits and direct R&D subsidies. My findings suggest that, as countries become more open to trade, they tend to switch from direct R&D subsidies to R&D tax credits. Since R&D tax credits tend to discriminate less between firms in the same country than direct subsidies, this pattern appears consistent with the model.

The findings of this chapter are potentially relevant in the context of the European Union. Over the past decades, the European Union has established the integrated "European Single Market" while government support for business R&D continues to be national. The present chapter may help facilitate a better understanding of the politico-economic considerations that are relevant in this context.

4 Tables and Figures

Chapter 1

Table 1: Cost of submitted projects and funding amounts received

	All funding applications (4730 total)		Baseline sample (2922 total)	
	Cost	Funding amount	Cost	Funding amount
# Project cost and *pwert* score available	4183	4183	2619	2619
# Projects approved for funding		2938		1735
Mean (in Euro)	662154	153682	529212	117984
Minimum	227	1500	227	2000
.25 Percentile	145950	39296	130000	34182
.5 Percentile	318000	78280	276935	67053
.75 Percentile	657950	156585	535450	128706
.9 Percentile	1404720	322983	1061092	238265
Maximum	26300000	3929000	17080100	2300000

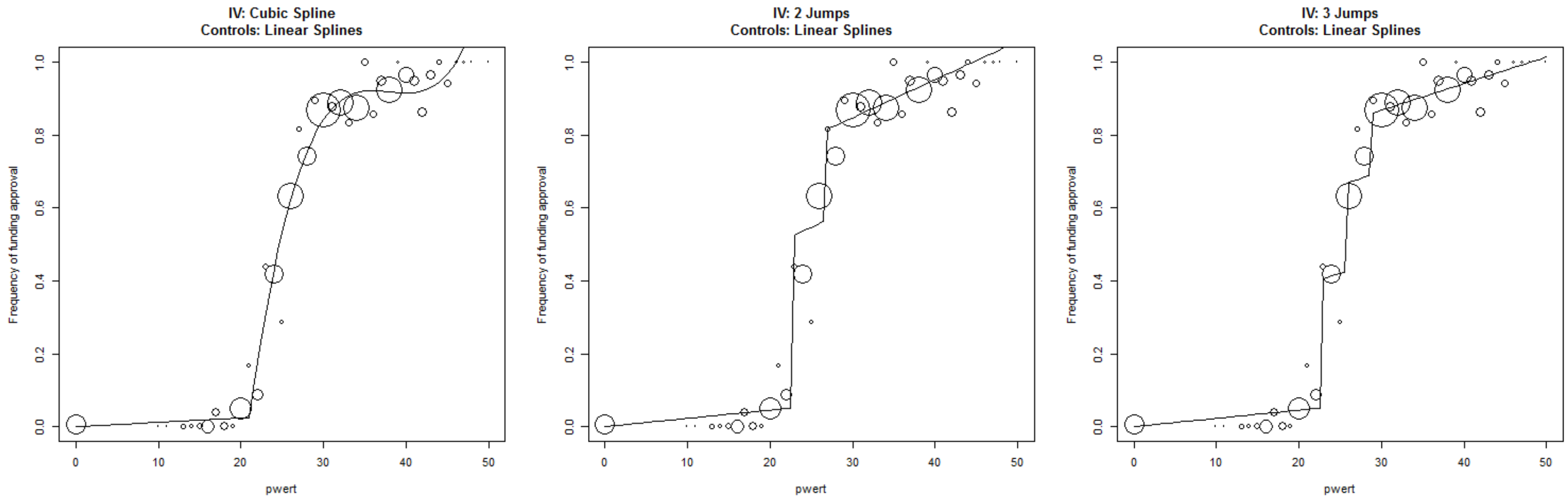
Note: Baseline sample is restricted to applications from firms that did not file multiple applications in the same year.

Table 2: Descriptive statistics for firms in the baseline sample

	R&D Exp in Euro	Employees	Sales in Euro	Age	# EP Patents 12 years before	# EP Patents 4 years after	# National Patents 12 years before	# National Patents 4 years after
# data available	1426	1508	1508	1791	1936	1936	1936	1936
Mean	1064602	114	30075480	11.13	1.56	1.05	2.45	0.91
Minimum	0	0	0	-12	0	0	0	0
.25 Percentile	30000	5	446250	1	0	0	0	0
.5 Percentile	150000	19	2463000	6	0	0	0	0
.75 Percentile	556500	86	4090000	15	1	1	1	1
.9 Percentile	1806500	277	54126400	30	3	2	4	2
Maximum	279216000	17000	7664000000	140	114	46	1067	101

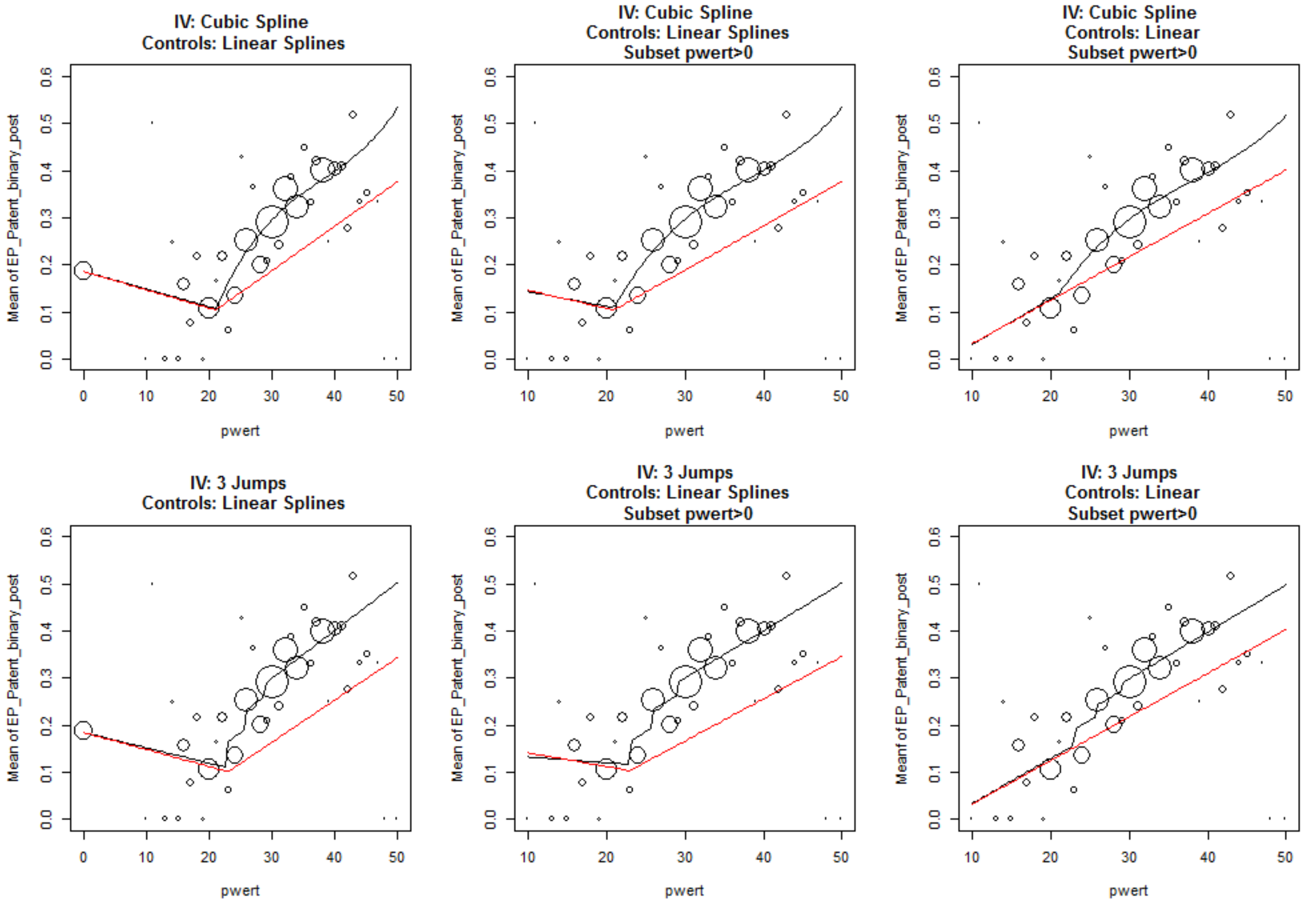
Note: Data for the first observed funding application per firm. Fraction of firms with at least one European Patent (National Patent) is 0.26 (0.27) in the 12 years before and 0.26 (0.23) in the 4 years after (and including) the year of the first observed funding application.

Figure 1: Fit for First Stage Regression: Frequency of funding approval as a function of linear splines and a non-linear increase in *pwert* score



Note: Areas of circles are proportional to number of observations. Three models: cubic spline (left), two discrete jumps (middle) and three discrete jumps (right) as instruments for the non-linear increase. See section 1.3 for a description of the models.

Figure 2: Fit for Second Stage Regression: Average propensity to file a European Patent as a function of linear splines in *pwert* score and non-linear increase in funding approval probability



Note: Areas of circles are proportional to number of observations. Two models: cubic spline (upper row) and three discrete jumps (lower row) as instruments for the non-linear increase in funding approval probability. Linear splines included as controls for *pwert* score. Left column shows the main specification, middle column excludes observeration with *pwert*=0 and right column controls linearly for score (discussed as robustness checks). See section 1.4 for further description.

Table 3: The effect of funding approval on the propensity to file a European Patent

	Dependent variable: EP Patent binary post					
	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 1	Model 2	Model 2	Model 3	Model 3
	IV: Cubic Spline	IV: Cubic Spline	IV: 2 jumps	IV: 2 jumps	IV: 3 jumps	IV: 3 jumps
approved	0.128**	0.108**	0.122**	0.101**	0.157***	0.111***
	(0.0561)	(0.0468)	(0.0522)	(0.0431)	(0.0517)	(0.0426)
pwert	-0.00390**	-0.00303*	-0.00275	-0.00258	-0.00362*	-0.00281*
	(0.00198)	(0.00164)	(0.00200)	(0.00165)	(0.00198)	(0.00163)
linear spline pwert 21	0.0133***	0.00405				
	(0.00340)	(0.00287)				
linear spline pwert 23			0.0131***	0.00413	0.0126***	0.00401
			(0.00315)	(0.00265)	(0.00317)	(0.00266)
project cost (in 100K Euro)		0.00306*		0.00305*		0.00306*
		(0.00172)		(0.00172)		(0.00172)
project cost sqr		-8.07e-06		-8.02e-06		-8.11e-06
		(1.29e-05)		(1.29e-05)		(1.29e-05)
EP Patent binary pre		0.180***		0.180***		0.180***
		(0.0327)		(0.0327)		(0.0327)
EP Patents pre		0.0238***		0.0238***		0.0238***
		(0.00300)		(0.00300)		(0.00300)
EP Patents pre sqr		-0.000259***		-0.000258***		-0.000259***
		(4.42e-05)		(4.43e-05)		(4.42e-05)
Nat Patent binary pre		0.136***		0.136***		0.136***
		(0.0291)		(0.0291)		(0.0291)
Nat Patents pre		0.00235*		0.00235*		0.00235*
		(0.00141)		(0.00141)		(0.00141)
Nat Patents pre sqr		-2.01e-06		-2.02e-06		-2.01e-06
		(1.24e-06)		(1.24e-06)		(1.24e-06)
Year FE, Sector FE, Age FE	NO	YES	NO	YES	NO	YES
Observations	2,619	2,619	2,619	2,619	2,619	2,619
R-squared	0.046	0.337	0.047	0.337	0.045	0.337

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Dependent variable is indicator that takes on value 1 if the firm filed an application for a European Patent in the year of funding application or during the subsequent three years. In Columns 1-6, I show results for alternative instruments for the increase in funding probability, illustrated in Figure 1 and Figure 2. Mean of dependent variable is 0.278. For a definition of control variables, see section 1.3.3.

Table 4: OLS estimates of effect of funding approval on the propensity to file a European Patent and Placebo regression on the 4 years preceding the funding application

	(1)	(2)	(3)	(4)
	EP Patent binary post	EP Patent binary post	EP Patent binary lagged 4y	EP Patent binary lagged 4y
	OLS	OLS	OLS	OLS
			Placebo Test	Placebo Test
approved	0.174***	0.0690***	0.133***	0.0423***
	(0.0182)	(0.0156)	(0.0176)	(0.0151)
Controls	NO	YES	NO	YES
Observations	2,619	2,619	2,619	2,619
R-squared	0.034	0.338	0.021	0.328

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. In models (2) and (4), same controls included in Table 3 for patenting histories, project cost, sector, age and year of application. In model (4), patent outcomes and controls are lagged by 4 years.

Table 5: Placebo regression for the effect of funding approval on the propensity to file a European Patent in the 4 years preceding the funding application

	Dependent variable: EP Patent binary for the 4 years preceding the funding application					
	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 1	Model 2	Model 2	Model 3	Model 3
	IV: Cubic Splines	IV: Cubic Splines	IV: 2 jumps	IV: 2 jumps	IV: 3 jumps	IV: 3 jumps
approved	0.0317	-0.00997	0.0370	-0.00546	0.0756	0.0155
	(0.0556)	(0.0464)	(0.0510)	(0.0425)	(0.0509)	(0.0418)
Controls	NO	YES	NO	YES	NO	YES
Observations	2,619	2,619	2,619	2,619	2,619	2,619
R-squared	0.034	0.330	0.034	0.330	0.035	0.330

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Same specifications as in Table 3, but patent outcomes and controls are lagged by 4 years.

Table 6: Heterogenous treatment effect of funding approval on the propensity to file a European Patent by firm age

	Dependent variable: EP Patent binary post					
	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
	IV: Cubic Splines	IV: 2 jumps	IV: 3 jumps	IV: Cubic Splines	IV: 2 jumps	IV: 3 jumps
	age$\leq 5y$	age$\leq 5y$	age$\leq 5y$	age$> 5y$	age$> 5y$	age$> 5y$
approved	0.0363	0.0386	0.0490	0.170**	0.151**	0.162***
	(0.0659)	(0.0608)	(0.0602)	(0.0663)	(0.0600)	(0.0596)
Observations	1,361	1,361	1,361	1,258	1,258	1,258
R-squared	0.298	0.299	0.299	0.314	0.317	0.315

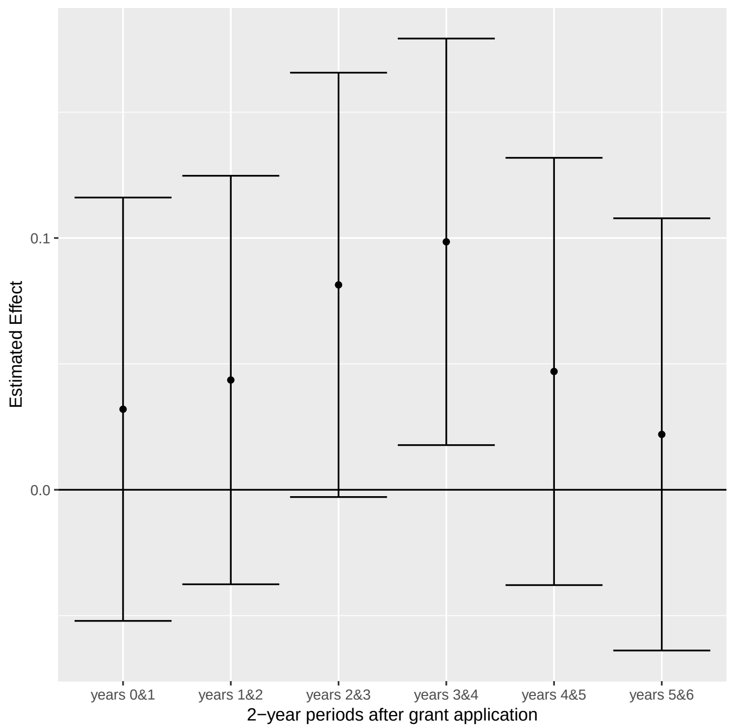
Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Full set of controls included in all models, but not reported. Mean of dependent variable is 0.322 for firms of age$>$ 5, 0.229 for firms age$\leq$ 5. Patent outcomes are measured for the year of funding application and the subsequent three years. Firms that are missing the year of incorporation counted as firms of age 0.

Table 7: The effect of funding approval on the propensity to file a patent by conventionality and quality for firms above the median age

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Unconventional Patent binary		Conventional Patent binary		High quality Patent binary		Low quality Patent binary	
	IV: Cubic Splines	IV: 3 jumps	IV: Cubic Splines	IV: 3 jumps	IV: Cubic Splines	IV: 3 jumps	IV: Cubic Splines	IV: 3 jumps
	age$> 5y$	age$> 5y$	age$> 5y$	age$> 5y$	age$> 5y$	age$> 5y$	age$> 5y$	age$> 5y$
approved	0.168**	0.174***	0.0263	0.0379	0.121*	0.138**	0.0382	0.0421
	(0.0688)	(0.0624)	(0.0667)	(0.0605)	(0.0665)	(0.0605)	(0.0686)	(0.0620)
Observations	1,258	1,258	1,258	1,258	1,258	1,258	1,258	1,258
R-squared	0.327	0.327	0.317	0.318	0.334	0.332	0.329	0.329

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Full set of controls included in all models, but not reported. See section 1.5.2 for a definition of variables. Mean of dependent variable is 0.35 for unconventional patents and 0.285 for conventional patents. Mean of dependent variable is 0.335 for high-quality patents and 0.297 for low-quality patents. Quality measured by number of citations received.

Figure 3: Time delay of the effect of funding approval on the Propensity to file a European Patent



Note: Figure shows point estimates of the treatment effect and 95% confidence intervals for consecutive two-year periods after funding application in percentage points (Year 0 is the year of the funding application).

Table 8: Descriptive evidence on utilization of technological knowledge novel to the firm for firms with patent applications before and after the funding application

	(1)	(2)	(3)
	avg_firm_novelty_post	avg_firm_novelty_post	avg_firm_novelty_post
	OLS	OLS	OLS
approved	0.200	0.302**	0.322**
	(0.151)	(0.139)	(0.137)
Controls	NO	NO	YES
Year FE, Sector FE, Age FE	NO	YES	YES
Observations	686	686	686
R-squared	0.003	0.059	0.145

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. $avg\_firm\_novelty$ is the average number of IPC classes cited per patent that are novel to the firm.

Chapter 2

Table 9: Summary statistics municipalities

	Mean	Std. Dev.	Min	25% Perc.	50% Perc.	75% Perc.	Max
All municipalities - 2580 obs							
# Foreign-Born Inventors (Residence)	2.14	19.45	0	0	0	1	675
# Domestic Inventors (Residence)	3.31	15.94	0	0	1	3	640
# Foreign-Born Inventors (Workplace)	2.11	27.97	0	0	0	0	1050
# Domestic Inventors (Workplace)	3.30	28.09	0	0	0	0	875
Tax rate income 200,000 CHF	19.09	2.66	6.16	17.53	19.55	20.89	25.56
Tax rate income 500,000 CHF	22.70	3.23	6.67	20.89	22.92	25.28	28.63
Municipality population size	3048.86	10703.95	12	466	1154	2855	372857
Crime rate	35.34	28.87	0	16.6	29.4	47.03	313.9
School graduation rate	19.32	6.54	7.14	14.29	18.01	22.78	42.17
Lakefront	0.11	0.31	0	0	0	0	1
Public transport usage	37.56	10.79	16.72	31.05	36.06	41.60	80.39
Distance national border (km)	21.36	16.75	0.01	6.79	18.08	32.61	68.54
Share German speakers	58.76	41.79	0	6.3	86.8	94.8	100
Share French speakers	27.36	40.23	0	0.4	1	82.13	100
State border municipalities - 782 obs							
# Foreign-Born Inventors (Residence)	1.25	3.28	0	0	0	1	38
# Domestic Inventors (Residence)	3.01	5.73	0	0	1	3	65
# Foreign-Born Inventors (Workplace)	0.88	4.85	0	0	0	0	67
# Domestic Inventors (Workplace)	2.46	12.18	0	0	0	0	217
Tax rate income 200,000 CHF	18.99	2.65	6.16	17.33	19.55	20.76	25.56
Tax rate income 500,000 CHF	22.45	3.30	6.67	20.72	22.67	24.86	28.48
Municipality population size	2622.54	4485.54	39	539	1235	3064	72959
Crime rate	34.21	26.97	0	15.93	28.2	46	240.4
School graduation rate	17.74	6.30	7.14	13.92	16.11	21.27	42.17
Lakefront	0.11	0.32	0	0	0	0	1
Public transport usage	36.23	9.61	16.72	30.34	35.27	41.14	62.11
Distance national border (km)	21.82	14.19	0.34	9.43	22.39	32.84	61.55
Share German speakers	70.66	37.58	0	38.75	91.15	95.5	100
Share French speakers	22.60	37.62	0	0.4	0.8	53.33	99.1

Note: Summary statistics for Swiss municipalities (as of 11-21-2010) for the year 2010. A municipality is included in the set of state border municipalities if it lies within 5 kilometers of distance of a municipality that belongs to a different state. See section 2.4 for definitions.

Table 10: Summary statistics for characteristics of inventor residence municipalities

	Mean	Std. Dev.	Min	25% Perc.	50% Perc.	75% Perc.	Max
All inventors - 14948 obs							
Commuting Distance	28.01	42.69	0	2.59	11.54	30.88	281.36
Tax rate income 200,000 CHF	18.17	2.80	6.16	16.35	18.73	20.08	25.56
Tax rate income 500,000 CHF	22.10	3.33	6.67	20.36	22.37	24.56	28.60
Municipality population size	62574.96	108866.3	37	3939	10812	51203	372857
Crime rate	73.38	43.64	0	37.3	61.8	110.6	283.5
School graduation rate	21.76	6.66	7.14	16.50	20.60	25.54	42.17
Lakefront	0.32	0.47	0	0	0	1	1
Public transport usage	50.62	16.57	16.72	36.97	47.10	59.72	80.39
Distance national border (km)	17.55	16.20	0.01	2.79	14.03	25.68	68.54
Share German speakers	64.34	35.68	0	11.88	82	88.3	99.4
Share French speakers	21.37	37.74	0	0.8	1.9	28.1	98.9
Inventors in state border municipalities - 3549 obs							
Commuting Distance	28.03	36.52	0	5.37	14.32	32.89	277.29
Tax rate income 200,000 CHF	17.65	3.16	6.16	16.26	18.20	19.87	25.56
Tax rate income 500,000 CHF	20.70	4.02	6.67	19.42	21.02	23.49	28.48
Municipality population size	10002.09	12746	108	2651	5468	12655	72959
Crime rate	49.32	26.56	0	30.4	45.3	64.7	169.1
School graduation rate	18.92	5.95	7.14	14.4	17.98	22.03	42.17
Lakefront	0.18	0.39	0	0	0	0	1
Public transport usage	42.32	9.82	16.72	34.57	41.66	48.57	62.11
Distance national border (km)	18.78	15.85	0.34	4.42	15	28.56	61.55
Share German speakers	77.75	27.45	0.6	83.1	87	91.2	99.3
Share French speakers	11.43	27.23	0	0.5	1	1.6	97.6

Note: Summary statistics for Swiss municipalities (as of 11-21-2010) for the year 2010. Data shown in this table is at the inventor level. A municipality is included in the set of state border municipalities if it lies within 5 kilometers of distance of a municipality that belongs to a different state.

Table 11: Main results on top income tax rates and the residential location choice of inventors

Dependent variable: choice of municipality of residence						
	Instrumented municipal top income tax rate (Average of the tax rates of all other municipalities in the same state that lie farther than 5km away)			Actual municipal top income tax rate		
	(1)	(2)	(3)	(4)	(5)	(6)
Subset of inventors	All	All	All	All	All	All
log net-of-tax rate	5.9437***	4.5718***	4.5282***	5.9733***	5.5096***	5.5110***
	[4.499, 7.306]	[2.857, 6.157]	[2.844, 6.164]	[4.859, 7.076]	[4.231, 7.116]	[3.988, 7.160]
Controls for municipal amenities						
Linear	NO	YES	YES	NO	YES	YES
Non-linear	NO	NO	YES	NO	NO	YES
Observations	12891	12891	12891	12891	12891	12891
avg. McFadden R-squared	0.0189	0.2437	0.2853	0.0273	0.2491	0.2873

Note: Maximum likelihood estimates from the residential location choice model. 95% confidence intervals reported, *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Log net-of-tax rate is given by $\log(1 - \text{tax rate}/100)$. All tax rates are average tax rates for an annual income of 500,000 CHF. Columns (2),(3),(5) and (6) include controls for commuting distance, municipality population size, lakefronts, distance to national border, crime rate, school graduation rate and public transport usage. In columns (3) and (6), I further include quadratic terms for all distance measures and cubic terms for municipality population size.

Table 12: Main results on top income tax rates: foreign-born vs domestic inventors

Dependent variable: choice of municipality of residence						
	Instrumented municipal top income tax rate (Average of the tax rates of all other municipalities in the same state that lie farther than 5km away)			Actual municipal top income tax rate		
	(1)	(2)	(3)	(4)	(5)	(6)
Subset of Inventors	Foreign-Born	Domestic	All	Foreign-Born	Domestic	All
log net-of-tax rate	5.9858***	4.1283***	3.9399***	6.4003***	5.3192***	5.1513***
	[3.5285, 8.0541]	[2.160, 6.031]	[2.047, 5.692]	[4.038, 9.412]	[3.816, 7.218]	[3.587, 7.0194]
log net-of-tax rate×Foreign-Born			2.4828			1.4426
			[-1.624, 7.508]			[-1.234, 5.453]
Controls for municipal amenities						
Linear	YES	YES	YES	YES	YES	YES
Observations	3825	9066	12891	3825	9066	12891
avg. McFadden R-squared	0.2922	0.2279	0.2440	0.2966	0.2337	0.2494

Note: Maximum likelihood estimates from the residential location choice model. 95% confidence intervals reported, *** p<0.01, ** p<0.05, * p<0.1. Log net-of-tax rate is given by $\log(1 - \text{tax rate}/100)$. All tax rates are average tax rates for an annual income of 500,000 CHF. All columns include controls for commuting distance, municipality population size, lakefronts, distance to national border, crime rate, school graduation rate and public transport usage. Foreign-born inventors are inventors who appear at least once with a nationality other than Swiss in the dataset (they may acquire Swiss citizenship over time). Domestic inventors are inventors that only appear as Swiss.

Table 13: Supporting evidence on top income tax rates and residential location choice of inventors: panel evidence on the effect of a reduction in top income tax rates in the states of Thurgau and Sankt Gallen on changes in residential location choices in state border municipalities over time

Dependent variable: choice of state border region of residence		
	Municipalities within 5 km of distance to the state border	Municipalities within 7.5 km of distance to the state border
	(1)	(2)
Subset of inventors	All	All
log net-of-tax rate	6.7265	6.9366
	$[-3.861, 18.572]$	$[-2.111, 17.134]$
Border region fixed effects	YES	YES
Observations	1576	2374
R-squared	0.2210	0.2028

Note: Maximum likelihood estimates from panel version of the residential location choice model described in section 2.5.2. 95% confidence intervals reported, *** p<0.01, ** p<0.05, * p<0.1. Evidence for inventors living around the state borders of the states of Thurgau and Sankt Gallen between 2005 and 2012. I record the first year in which an inventor appears in a state border region. In column (1), municipalities within 5 km of distance to a municipality that belongs to a different state are considered. In column (2), the maximal distance is expanded to 7.5 km. The p-values of the coefficients in column (1) and (2) are 0.24 and 0.16 respectively (two-sided test). Analysis is carried out at the level of state border regions (see section 2.5.2 for definition).

Table 14: Municipal amenities and the residential location choice of inventors

Dependent variable: choice of municipality of residence				
	Actual municipal top income tax rate			
	(1)	(2)	(3)	(4)
Subset of Inventors	All	All	≤ 80km comm. dist.	≤ 80km comm. dist.
tax rate (pp)	−0.0756***	−0.0515***	−0.0793***	−0.0620***
	[-0.095, -0.056]	[-0.074, -0.026]	[-0.100, -0.058]	[-0.083, -0.036]
commuting distance (km)	−0.1829***	−0.2543***	−0.2001***	−0.2900***
	[-0.209, -0.163]	[-0.288, -0.229]	[-0.228, -0.179]	[-0.330, -0.259]
commuting distance squared		0.0014***		0.0023***
		[0.001, 0.002]		[0.002, 0.003]
population size (thousd.)	0.0906***	0.2449***	0.0930***	0.2187***
	[0.073, 0.109]	[0.214, 0.282]	[0.078, 0.112]	[0.196, 0.258]
population size squared		−0.0069***		−0.0057***
		[-0.008, -0.005]		[-0.007, -0.004]
population size cubic		5.83e-05***		4.63e-05***
		[4.1e-05,8.0e-05]		[3.2e-05,6.6e-05]
lakefront (binary)	0.3967***	0.3579***	0.4268***	0.4197***
	[0.221, 0.639]	[0.172, 0.605]	[0.234, 0.695]	[0.214, 0.674]
distance national border (km)	−0.0295**	−0.0523***	−0.0212	−0.0486
	[-0.057, -0.004]	[-0.091, -0.010]	[-0.050, 0.004]	
distance national border squared		0.0008**		0.0007
		[7.6e-05, 0.0013]		[-0.0002, 0.001]
Other controls:				
school grad. rate, crime rate, public transport usage	YES	YES	YES	YES
Observations	12891	12891	11993	11993
avg. McFadden R-squared	0.2490	0.2870	0.2593	0.2910

Note: Maximum likelihood estimates from the residential location choice model. 95% confidence intervals reported, *** p<0.01, ** p<0.05, * p<0.1. All tax rates are average tax rates for an annual income of 500,000 CHF. I present results for the specification that includes the tax rate (as opposed to the log net-of-tax rate) to facilitate the comparison of the coefficients. In columns (3) and (4), I restrict sample to inventors who reside no further than 80 kilometers away from their employer.

Table 15: Sorting by language: relationship between residential location choice and language spoken in the municipality

Dependent variable: choice of municipality of residence						
	(1)	(2)	(3)	(4)	(5)	(6)
Subset of Inventors	All	All	Foreign-Born	Foreign-Born	Domestic	Domestic
share german speakers (pp)	−0.0025		−0.0121		−0.0017	
	[-0.009, 0.004]		[-0.042, 0.007]		[-0.008, 0.004]	
share german speakers ×german sp inventor (pp)			0.0112			
			[−0.015, 0.048]			
share french speakers (pp)		0.0035		0.0028		0.0039
		[-0.003, 0.010]		[-0.014, 0.018]		[-0.002, 0.010]
share french speakers ×french inventor (pp)				−0.0029		
				[-0.094, 0.031]		
Controls for municipal amenities and tax rates	YES	YES	YES	YES	YES	YES
Observations	12891	12891	3825	3825	9066	9066
avg. McFadden R-squared	0.2912	0.2915	0.3602	0.3593	0.2700	0.2702

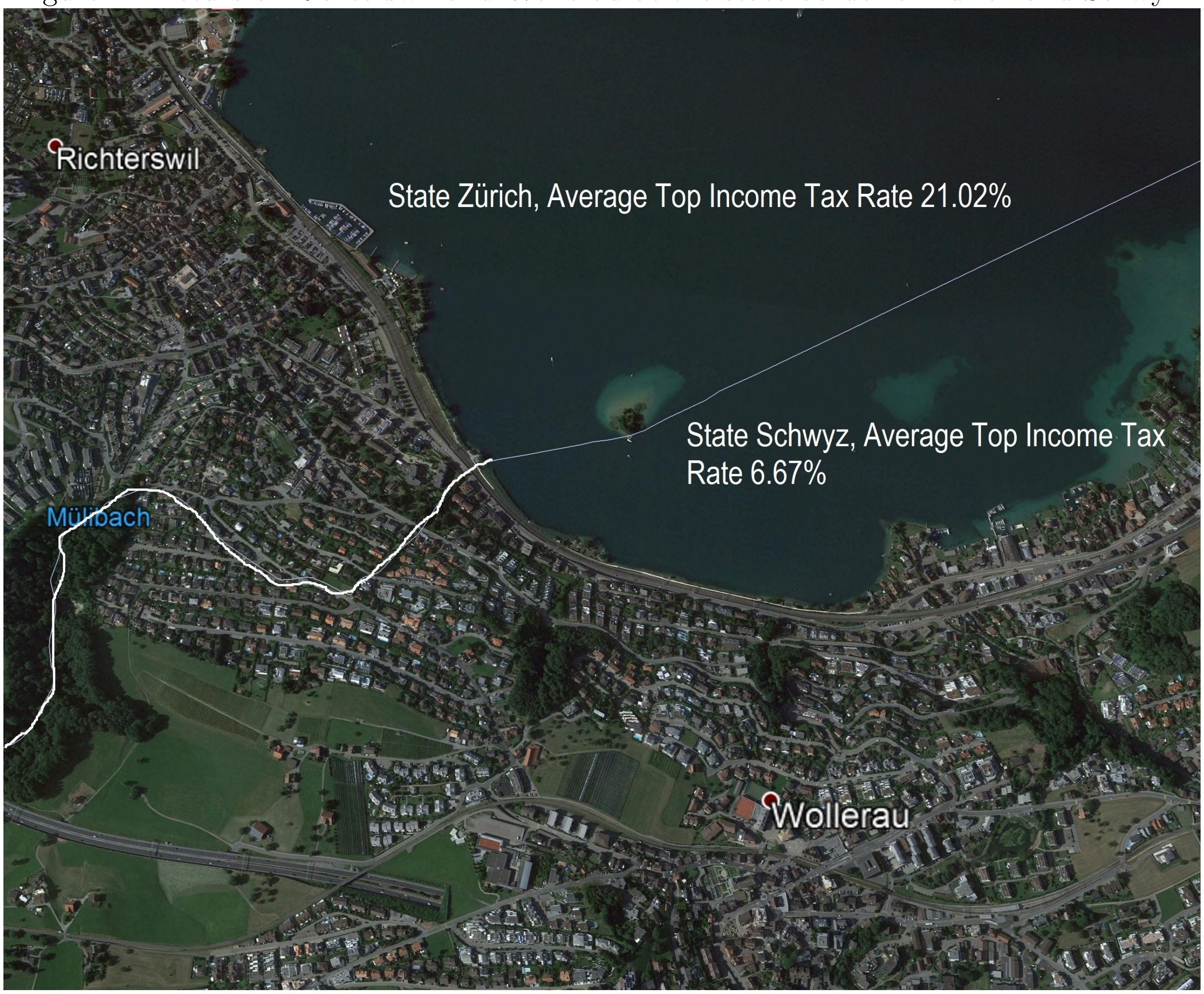
Note: Maximum likelihood estimates from the residential location choice model. 95% confidence intervals reported, *** p<0.01, ** p<0.05, * p<0.1. All columns include controls for tax rate, commuting distance, municipality population size, lakefronts, distance to national border, crime rate, school graduation rate and public transport usage. Foreign-born inventors are inventors who appear at least once with a nationality other than Swiss in the dataset. Domestic inventors are inventors that only appear as Swiss. The variable german sp inventor is an indicator that takes on 1 if the inventor is from either Germany or Austria. Share german speakers is the share of the population in the municipality that named German as their primary language. Variables for french speakers are defined analogously.

Table 16: Sorting by political attitudes: residential location choice and attitude towards immigration in the municipality

Dependent variable: choice of municipality of residence						
	(1)	(2)	(3)	(4)	(5)	(6)
Subset of Inventors	All	All	Foreign-Born	Foreign-Born	Domestic	Domestic
vote share anti-immigration party (pp)	-0.0145^{***}	-0.076	-0.0054	0.0058	-0.0181^{***}	-0.0118^{*}
	[-0.022, -0.004]	[-0.015, 0.001]	[-0.028, 0.014]	[-0.013, 0.026]	[-0.027, -0.007]	[-0.021, 0.0002]
share german speakers (pp)		-0.0043		-0.0131		-0.0032
		[-0.011, 0.002]		[-0.033, 0.004]		[-0.010, 0.004]
average income (thousd.)		-0.0037		-0.0048		-0.0039
		[-0.009, 0.003]		[-0.016, 0.007]		[-0.010, 0.002]
social status index		0.0427^{***}		0.0529^{***}		0.0447^{***}
		[0.028, 0.061]		[0.026, 0.087]		[0.025, 0.067]
Controls for municipal amenities and tax rates	YES	YES	YES	YES	YES	YES
Observations	12891	12891	3825	3825	9066	9066
avg. McFadden R-squared	0.2927	0.2980	0.3566	0.3649	0.2731	0.2789

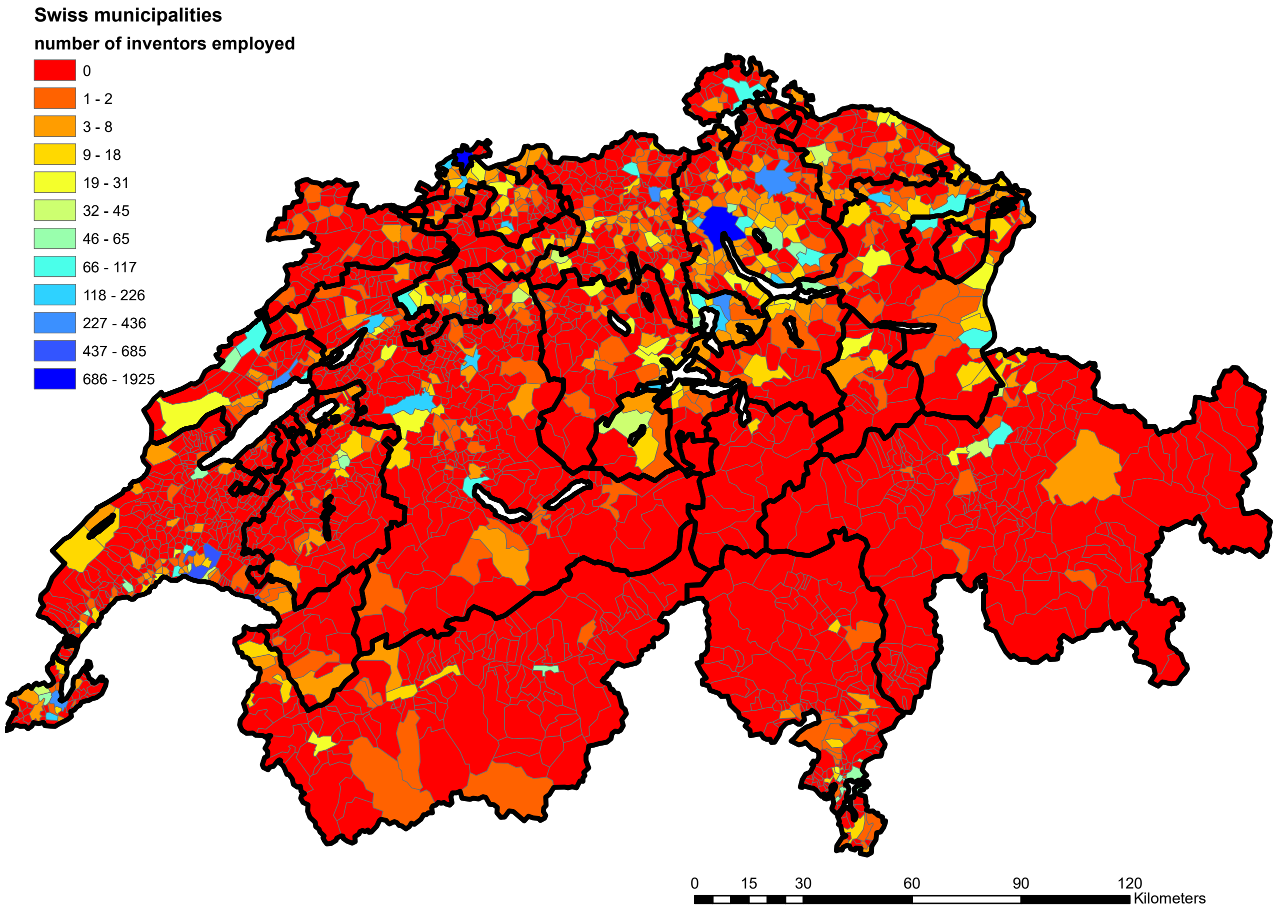
Note: Maximum likelihood estimates from the residential location choice model. 95% confidence intervals reported, *** p<0.01, ** p<0.05, * p<0.1. All columns include controls for tax rate, commuting distance, municipality population size, lakefronts, distance to national border, crime rate, school graduation rate and public transport usage. Foreign-born inventors are inventors who appear at least once with a nationality other than Swiss in the dataset. Domestic inventors are inventors that only appear as Swiss. The variable vote share anti-immigration party is the vote share of the SVP ("Schweizerische Volkspartei") in the municipality in the federal elections of 2011. The variable social status index is a measure by the Swiss Statistical Office based on the share of university graduates, high income earners and individuals in managing positions living in the municipality.

Figure 4: Picture of Richterswil and Wollerau at the state border of Zürich and Schwyz.



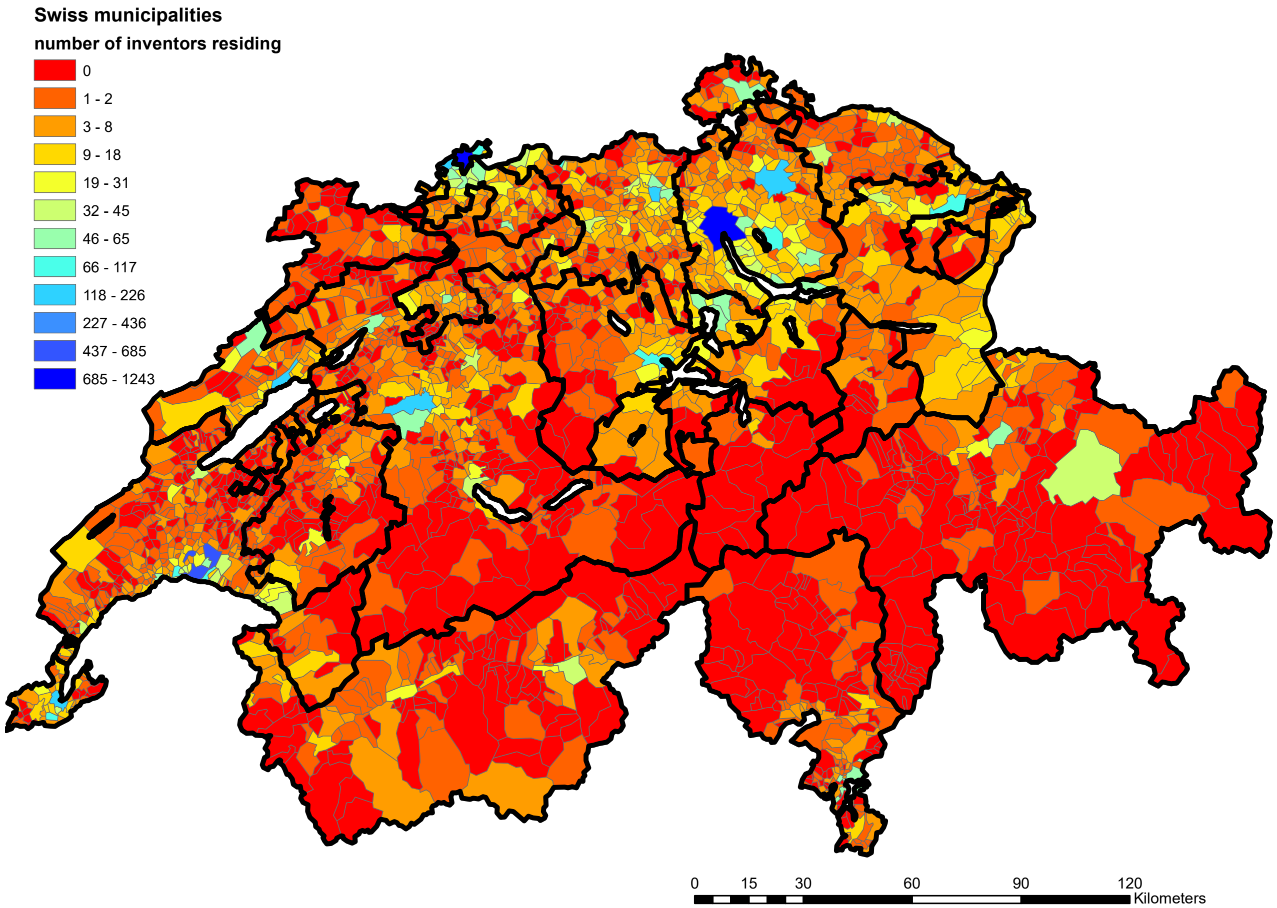
Note: Picture from Google Earth. The average income tax rate is 6.67 in Wollerau and 21.02 in Richterswil for an individual earning 500,000 CHF (around 450,000 USD) annually in 2010. Tax rate is tax burden in percent of the gross income levied by the state and the municipality, excluding the federal income tax, for a single income earner with no children earning 500,000 CHF (around 450,000 USD) in 2010.

Figure 5: Distribution inventor workplaces



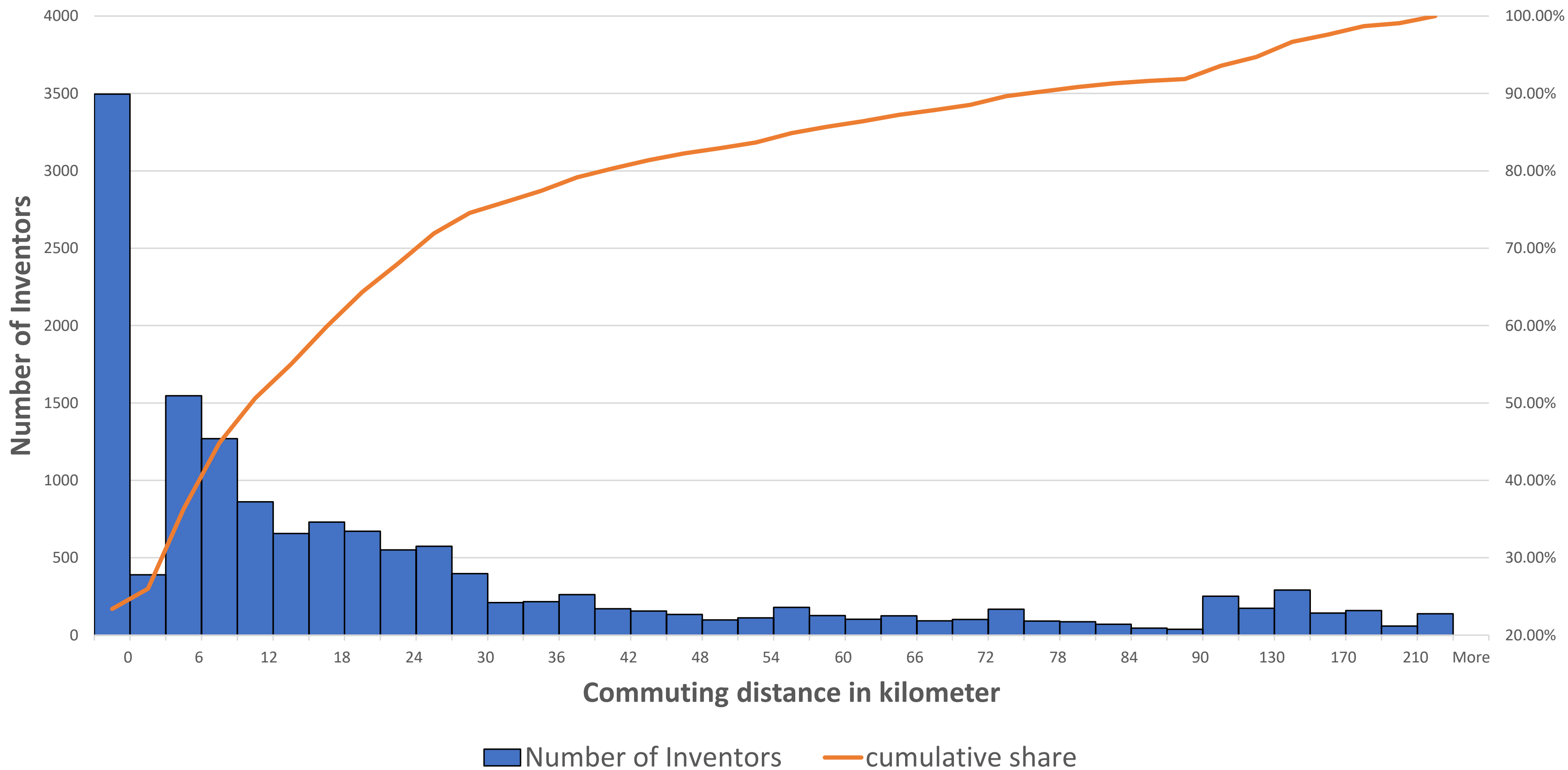
Note: Swiss municipalities as of 11-21-2010. Included inventors are inventors listed on at least one PCT patent filed between 2005 and 2012. State borders appear as thick lines, municipality borders appear as thin lines.

Figure 6: Distribution inventor residence locations



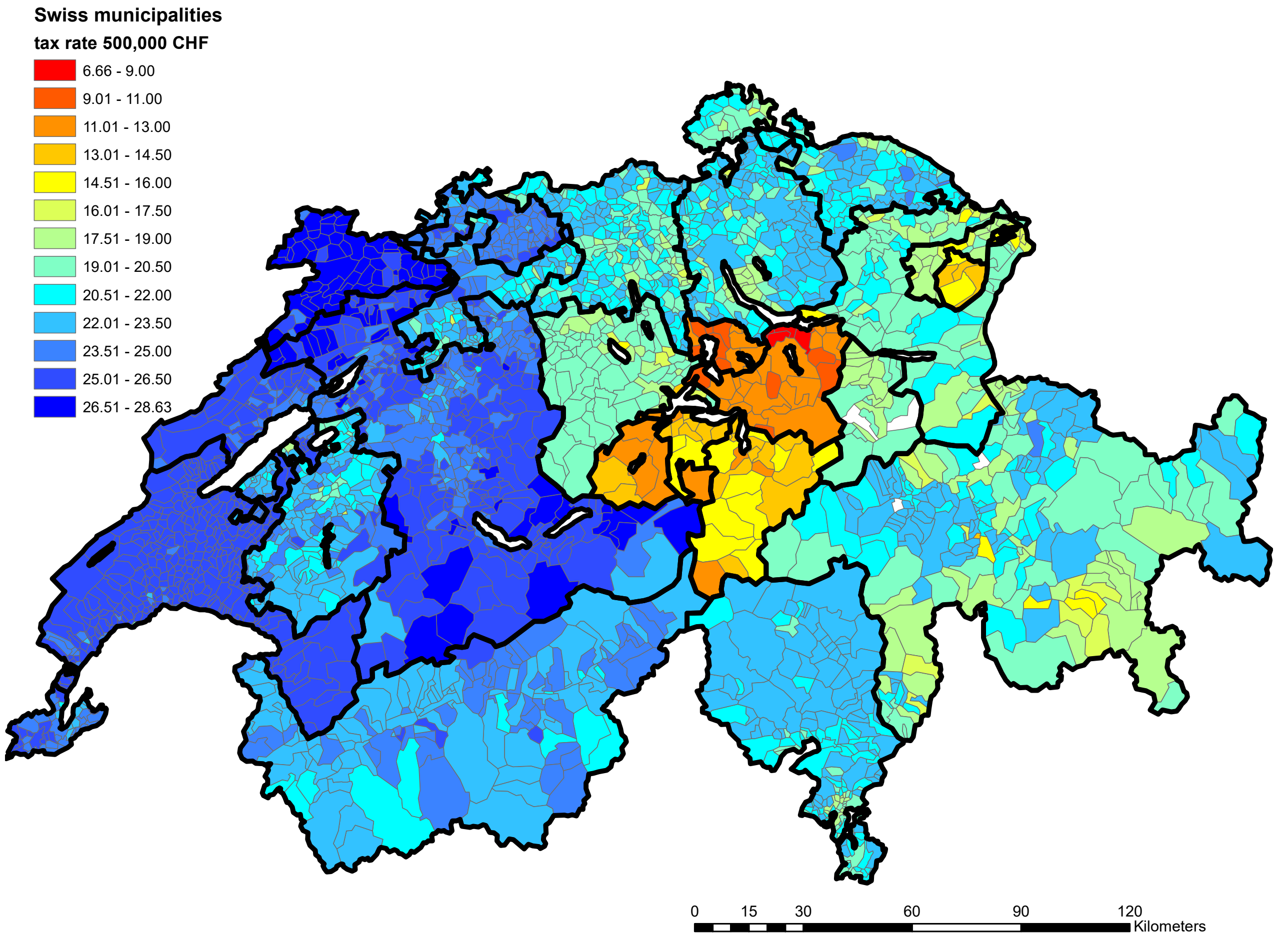
Note: Swiss municipalities as of 11-21-2010. Included inventors are inventors listed on at least one PCT patent filed between 2005 and 2012. State borders appear as thick lines, municipality borders appear as thin lines.

Figure 7: Distribution of commuting distance



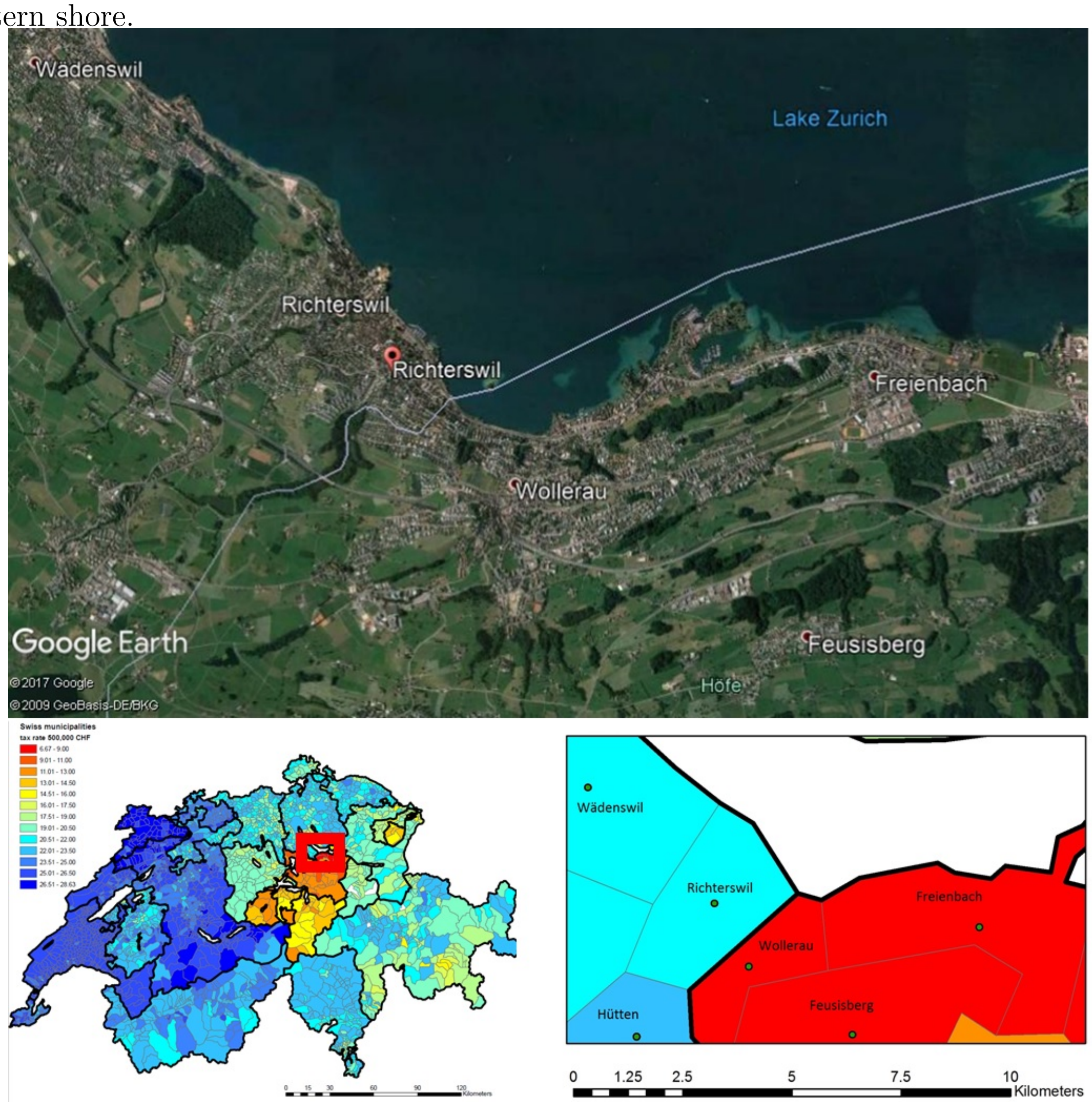
Note: Commuting distance is distance between centroid of inventor residence municipality and centroid of inventor workplace municipality. Distance equal 0 (leftmost bar) means that residence and workplace lie in the same municipality.

Figure 8: Tax burden in percent of the gross income levied by the state and the municipality, excluding the federal income tax, for an unmarried individual with no children earning 500,000 CHF annually.



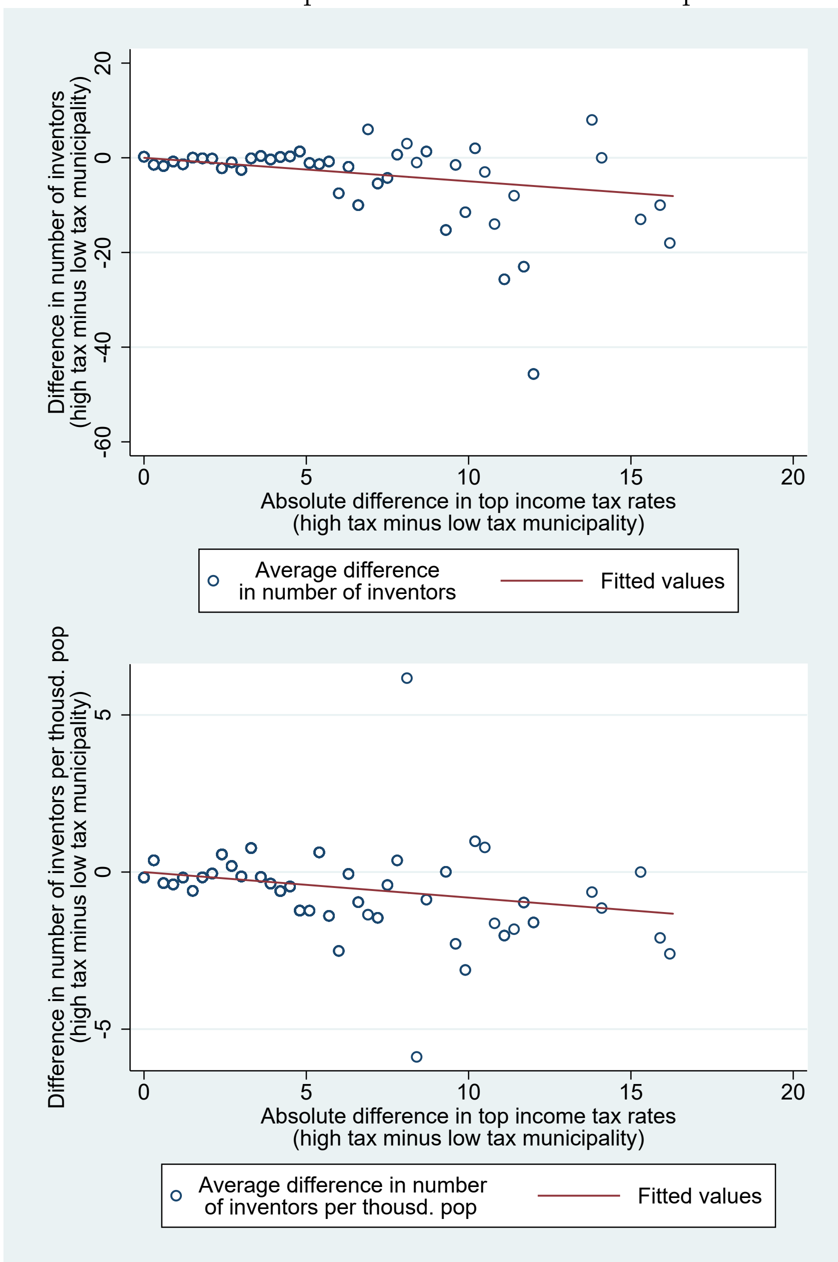
Note: Swiss municipalities as of 11-21-2010. Tax rates as of 01-01-2010. Tax rates missing for 4 out of 2584 municipalities (Matt, Luchsingen, Felsberg, Mundaun). State borders appear as thick lines, municipality borders appear as thin lines.

Figure 9: Border region between states of Zürich and Schwyz. Municipalities with respective top income tax rates are: Richterswil 21.02, Wädenswil 21.56 (Zürich) and Wollerau 6.67, Feusisberg 6.95, Freienbach 6.89 (Schwyz). City of Zürich is located further North on the Western shore.



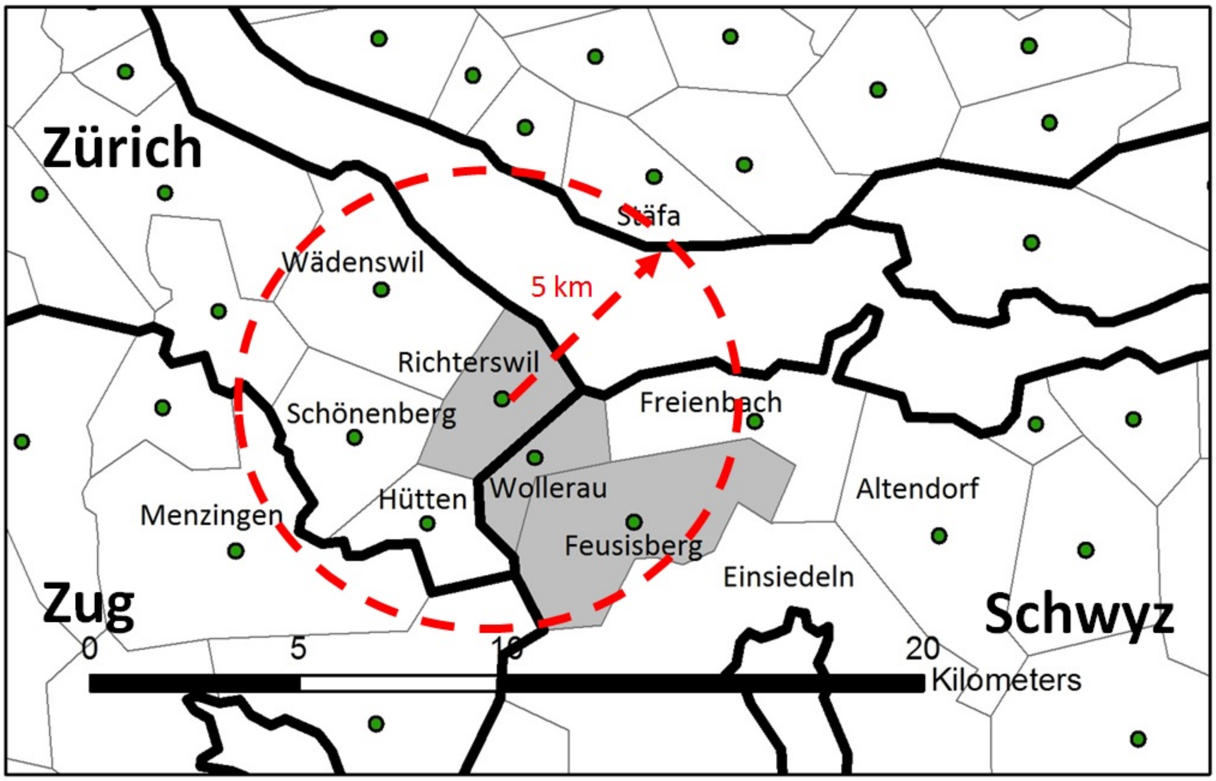
Note: Swiss municipalities as of 11-21-2010. Tax rates as of 01-01-2010. Top income tax rate is tax burden in percent of the gross income levied by the state and the municipality, excluding the federal income tax, for an unmarried individual with no children earning 500,000 CHF annually.

Figure 10: Relationship between the difference in top income tax rates and the difference in the number of resident inventors for pairs of state border municipalities



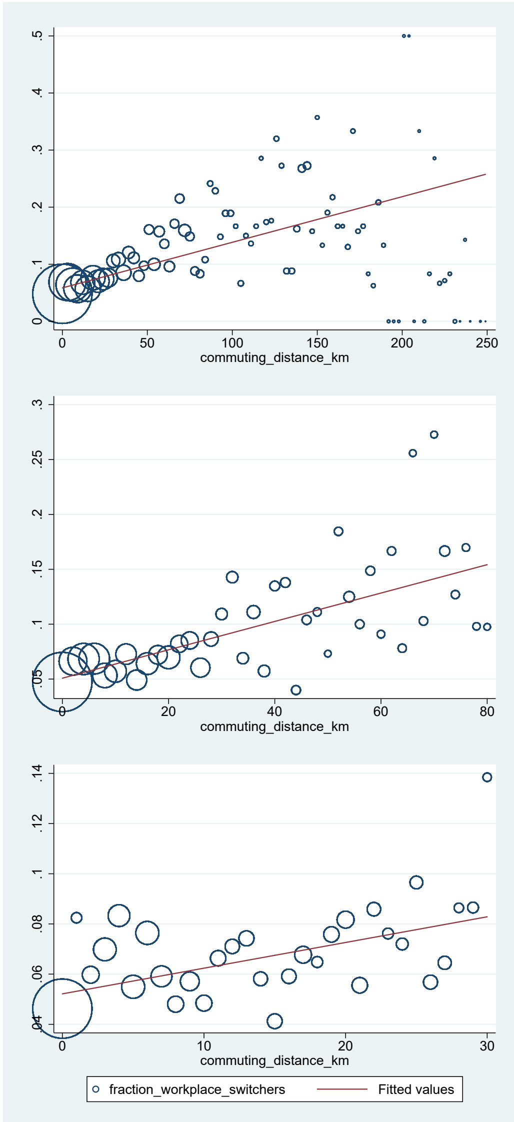
Note: Plots show the relationship between the absolute difference in tax rates and the difference in the number of resident inventors (inventors in high-tax municipality minus inventors in low-tax municipality, left graph) for all distinct pairs of municipalities that lie no farther than 5 kilometers apart and on different sides of state borders (number of inventors per thousand population in right graph). Averages are presented using a bin-size of 0.3 in absolute tax difference. Estimate of the slope coefficient is $-0.49^{***}(0.14)$ in the left graph and $-0.081^{**}(0.032)$ in the right graph from a regression without constant (standard errors clustered at the municipality level).

Figure 11: Example of inventor's choice set: picture shows the municipalities in the choice sets assigned to an inventor who resides in Richterswil at the state border between Zürich and Schwyz (colored in grey).



Note: Swiss municipalities as of 11-21-2010. In this example, the admissible choice pairs are (Richterswil,Wollerau) and (Richterswil,Feusisberg). Only municipalities that lie within 5 kilometers of distance to the municipality of residence and that belong to a different state are considered (besides the municipality of residence). Distance between municipalities is measured as straight-line distance between the centroids.

Figure 12: Commuting distance to initial workplace and the probability of switching workplace



Note: I restrict sample to inventors that reside within 30 km, 80 km or 250 km of distance to the initial employer. Averages for bin-sizes 1 km, 2 km and 3 km respectively are shown. Circle size is proportional to the number of inventors. Dependent variable takes on value 1 if change in the workplace of the inventor was observed, 0 otherwise. The workplace is defined as the municipality where the employer of the inventor is located. Thus, a switch between employers that are located at the same location does not constitute a change in workplace. Because the workplace is only observed when applying for a patent, the actual number of inventors who switched workplace might in reality be higher.

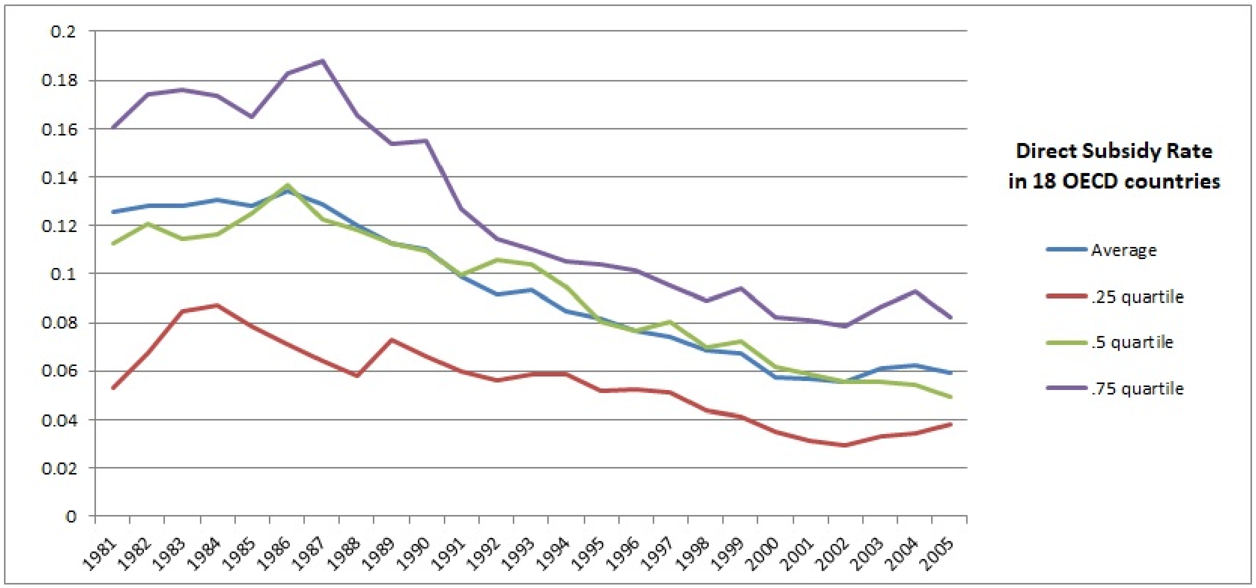
Table 17: Relationship between commuting distance to initial employer and the probability of switching workplace

Dependent variable: Inventor switched workplace (binary)			
	(1)	(2)	(3)
Subset of Inventors	≤ 250km comm. dist.	≤ 80km comm. dist.	≤ 30km comm. dist.
Model	OLS-LPM	OLS-LPM	OLS-LPM
commuting distance (to initial workplace in km)	0.000799***	0.00130***	0.00102***
	(5.76e-05)	(0.000125)	(0.000275)
Constant	0.0580***	0.0503***	0.0522***
	(0.00281)	(0.00306)	(0.00335)
Observations	12,772	11,682	9,753
R-squared	0.015	0.009	0.001

Note: Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Dependent variable takes on value 1 if change in the workplace of the inventor was observed, 0 otherwise. The workplace is defined as the municipality where the employer of the inventor is located. Thus, a switch between employers that are located at the same location does not constitute a change in workplace. Because the workplace is only observed when applying for a patent, the actual number of inventors who switched workplace might in reality be higher.

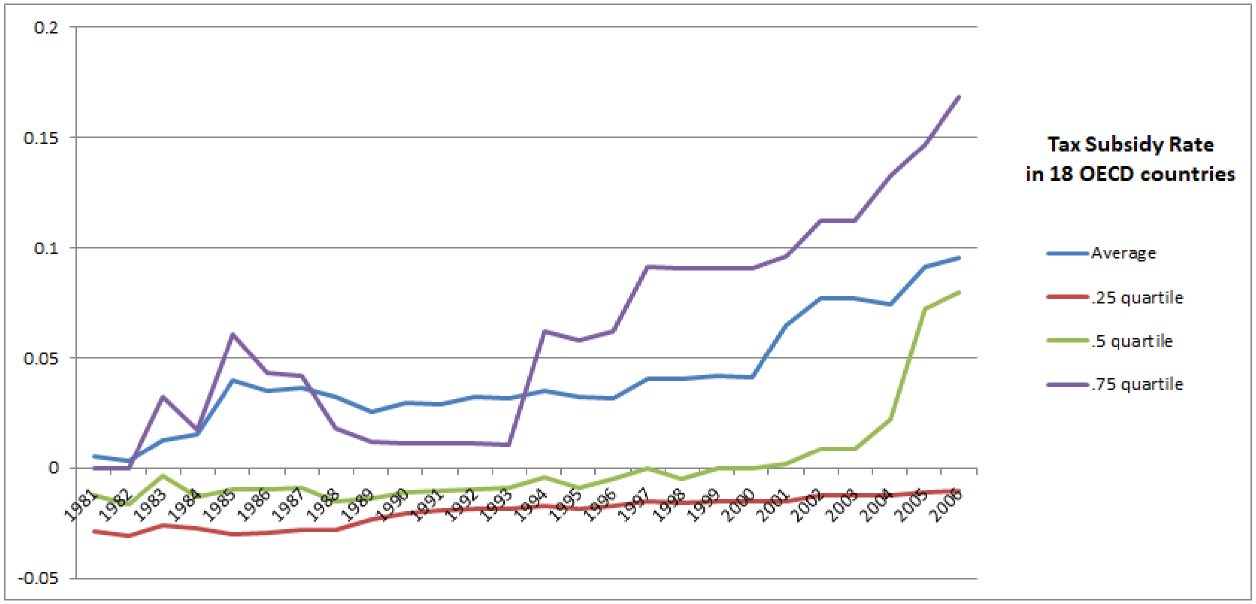
Chapter 3

Figure 13: Direct research subsidies for business research and development in the OECD



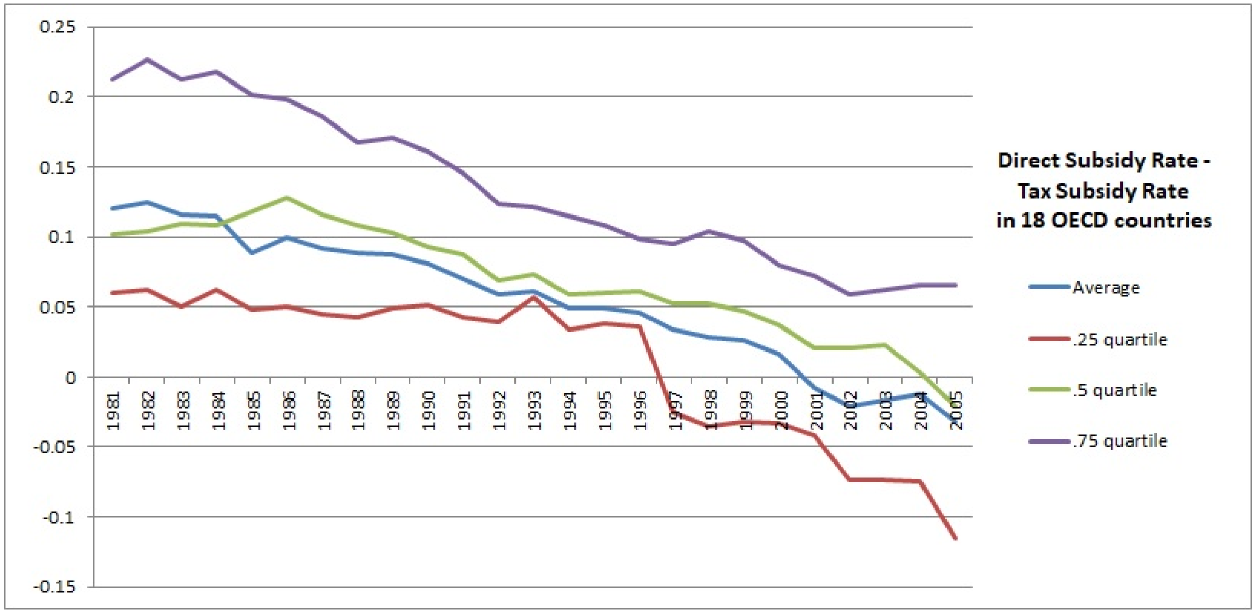
Graph shows share of business expenditures on R&D funded directly by government (the "direct subsidy rate") for 18 OECD countries from 1980 to 2006. Data source: OECD. Average, first, second and third quartile are shown. Countries included: Australia, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, UK, US.

Figure 14: R&D tax credits for business research and development in the OECD



Graph shows the tax subsidy rate implied by R&D tax credits for business expenditures on R&D in 18 OECD countries from 1980 to 2006. See definition in Section 3.3. Average, first, second and third quartile are shown. Data source: OECD. Countries included: Australia, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, UK, US.

Figure 15: Relative generosity of direct government funding and R&D tax credits for business R&D in the OECD



Graph shows the the difference between the share of business expenditures on R&D funded directly by government ("direct subsidy rate") and the tax subsidy rate implied by R&D tax credits for business expenditures on R&D in 18 OECD countries from 1980 to 2006. A decline indicates relatively more subsidization through R&D tax credits. See definition in Section 3.3. Average, first, second and third quartile are shown. Data source: OECD. Countries included: Australia, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, UK, US.

Table 18: Relationship between changes in trade openness and changes in government support for business R&D in 24 OECD countries 1980-2006

	Dependent variable:		
	ΔBERD (pp)	ΔSR	ΔRSR
	(1)	(2)	(3)
ΔOPEN (pp)	0.000315	0.000685	−0.001470
	(0.000697)	(0.000608)	(0.000597)**
	((0.000558))	((0.000596))	((0.000667))**
Year Dummy Variables	Yes	Yes	Yes
Observations	515	515	515
R^2	0.128	0.053	0.056

Note: *p<0.1; **p<0.05; ***p<0.01. Clustered standard errors (country-level) in double brackets. *Open* is trade openness in percentage points. *BERD* is business expenditure on R&D as share of GDP in percentage points. *SR* is a measure of the total generosity of government funding for business R&D. *RSR* is a measure of the relative generosity of R&D tax credits and direct funding, where a decrease indicates relatively more subsidization through tax credits.

Table 19: Simulation Results: Difference in government welfare between models with separated and integrated product markets in percent

$(\lambda-1)*100$ \ $w*100$	0.00	10.00	20.00	30.00	40.00	50.00	60.00	70.00	80.00	90.00	100.00	110.00	120.00	130.00	140.00	150.00	160.00	170.00	180.00
1.00	-0.05	-0.84	-1.44	-1.89	-2.25	-2.53	-2.75	-2.93	-3.07	-3.19	-3.28	-3.36	-3.42	-3.47	-3.51	-3.54	-3.56	-3.58	-3.60
2.00	-0.09	-0.88	-1.46	-1.91	-2.25	-2.52	-2.73	-2.90	-3.03	-3.14	-3.22	-3.29	-3.35	-3.39	-3.42	-3.44	-3.46	-3.47	-3.48
3.00	-0.13	-0.91	-1.48	-1.91	-2.24	-2.50	-2.70	-2.86	-2.98	-3.08	-3.15	-3.21	-3.26	-3.29	-3.32	-3.33	-3.34	-3.42	-3.66
4.00	-0.17	-0.93	-1.49	-1.91	-2.23	-2.48	-2.66	-2.81	-2.92	-3.01	-3.07	-3.12	-3.15	-3.18	-3.20	-3.37	-3.64	-3.88	-4.09
5.00	-0.20	-0.95	-1.50	-1.91	-2.21	-2.44	-2.61	-2.74	-2.84	-2.92	-2.97	-3.01	-3.03	-3.19	-3.48	-3.73	-3.95	-4.14	-4.31
6.00	-0.23	-0.97	-1.50	-1.89	-2.18	-2.39	-2.55	-2.66	-2.75	-2.81	-2.84	-2.87	-3.20	-3.48	-3.71	-3.90	-4.06	-4.19	-4.29
7.00	-0.25	-0.98	-1.50	-1.87	-2.14	-2.33	-2.47	-2.56	-2.63	-2.67	-2.81	-3.12	-3.38	-3.58	-3.73	-3.84	-3.91	-3.95	-3.95
8.00	-0.27	-0.98	-1.48	-1.83	-2.07	-2.24	-2.36	-2.43	-2.47	-2.66	-2.96	-3.18	-3.33	-3.43	-3.46	-3.45	-3.39	-3.29	-3.15
9.00	-0.28	-0.97	-1.45	-1.77	-1.99	-2.13	-2.22	-2.26	-2.45	-2.72	-2.89	-2.98	-2.98	-2.91	-2.76	-2.55	-2.27	-1.93	-1.51
10.00	-0.28	-0.95	-1.40	-1.69	-1.87	-1.98	-2.03	-2.17	-2.41	-2.52	-2.49	-2.36	-2.11	-1.75	-1.27	-0.68	0.05	0.92	1.96
11.00	-0.27	-0.92	-1.33	-1.58	-1.72	-1.77	-1.84	-2.03	-2.04	-1.86	-1.50	-0.95	-0.20	0.79	2.07	3.70	5.83	8.70	12.85
12.00	-0.25	-0.86	-1.22	-1.42	-1.49	-1.48	-1.60	-1.46	-1.03	-0.28	0.84	2.44	4.73	8.15	13.94	27.78	-356.75	-343.55	-324.97

Note: Table shows the difference in aggregate government welfare between the model with separated product markets and integrated product markets in percent of government welfare under separated product markets (definitions 2 and 4). A negative number indicates that aggregate government welfare is higher under separated product markets. Descriptions of the models can be found in Section 3.2. The parameter $(\lambda-1)*100$ (rows) measures the increase in quality in each innovative step in percent and the parameter $w*100$ (columns) measures the external benefit of research that accrues to the government in percent of the private benefit from innovation.

Table 20: Simulation Results: Difference in aggregate R&D spending between models with separated and integrated product markets in percent

$(\lambda-1)*100$ \ $w*100$	0.00	10.00	20.00	30.00	40.00	50.00	60.00	70.00	80.00	90.00	100.00	110.00	120.00	130.00	140.00	150.00	160.00	170.00
1.00	0.23	4.28	8.32	12.34	16.33	20.28	24.18	28.03	31.83	35.58	39.28	42.94	46.54	50.09	53.60	57.06	60.48	63.85
2.00	0.49	4.60	8.69	12.75	16.77	20.75	24.67	28.54	32.36	36.11	39.82	43.46	47.05	50.58	54.06	57.48	60.85	64.17
3.00	0.80	4.95	9.09	13.19	17.24	21.24	25.17	29.04	32.85	36.58	40.26	43.86	47.39	50.86	54.27	57.60	60.88	65.14
4.00	1.14	5.35	9.53	13.66	17.73	21.73	25.66	29.52	33.29	36.98	40.60	44.13	47.57	50.94	54.22	60.15	67.95	75.76
5.00	1.52	5.78	10.00	14.15	18.23	22.23	26.15	29.96	33.68	37.31	40.83	44.26	47.58	53.11	61.26	69.40	77.54	85.65
6.00	1.93	6.24	10.50	14.67	18.76	22.74	26.62	30.38	34.02	37.55	40.96	44.34	52.82	61.29	69.74	78.16	86.54	94.86
7.00	2.38	6.74	11.03	15.22	19.30	23.25	27.07	30.76	34.30	37.70	42.84	51.66	60.45	69.19	77.88	86.50	95.03	103.47
8.00	2.87	7.28	11.60	15.79	19.85	23.76	27.51	31.10	34.52	40.60	49.76	58.88	67.93	76.91	85.80	94.57	103.24	111.78
9.00	3.39	7.85	12.19	16.39	20.42	24.27	27.93	31.40	37.70	47.24	56.72	66.14	75.47	84.70	93.82	102.82	111.71	120.49
10.00	3.95	8.46	12.83	17.01	21.00	24.78	28.34	34.19	44.15	54.08	63.95	73.76	83.49	93.16	102.77	112.37	121.97	131.65
11.00	4.55	9.12	13.50	17.67	21.61	25.31	30.08	40.56	51.04	61.52	72.02	82.56	93.22	104.08	115.31	127.17	140.08	154.85
12.00	5.20	9.83	14.23	18.39	22.27	25.87	36.47	47.66	59.01	70.60	82.62	95.38	109.48	126.21	149.21	195.77	349.70	327.37 s

Note: Table shows the difference in aggregate R&D spending between the model with separated product markets and integrated product markets in percent of aggregate R&D spending under separated product markets (definitions 2 and 4). A positive number indicates that aggregate research spending is higher under integrated product markets. Descriptions of the models can be found in Section 3.2. The parameter $(\lambda-1)*100$ (rows) measures the increase in quality in each innovative step in percent and the parameter $w*100$ (columns) measures the external benefit of research that accrues to the government in percent of the private benefit from innovation.

Table 21: Simulation Results: R&D subsidies as a share of R&D spending under separated product markets

$(\lambda-1)*100$ \ $w*100$	0.00	10.00	20.00	30.00	40.00	50.00	60.00	70.00	80.00	90.00	100.00	110.00	120.00	130.00	140.00	150.00	160.00	170.00	180.00
1.00	0.01	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.09	0.10	0.12	0.13	0.15	0.17	0.19	0.21	0.23	0.25	0.28
2.00	0.03	0.06	0.09	0.12	0.16	0.20	0.24	0.29	0.35	0.41	0.47	0.53	0.61	0.68	0.76	0.85	0.93	1.03	1.12
3.00	0.08	0.13	0.20	0.27	0.36	0.45	0.55	0.67	0.79	0.92	1.07	1.22	1.38	1.55	1.73	1.92	2.11	2.32	2.53
4.00	0.14	0.24	0.36	0.49	0.64	0.81	1.00	1.20	1.42	1.66	1.92	2.19	2.47	2.77	3.09	3.42	3.77	4.13	4.51
5.00	0.23	0.39	0.57	0.78	1.02	1.29	1.59	1.91	2.26	2.63	3.03	3.45	3.90	4.37	4.86	5.37	5.91	6.46	7.04
6.00	0.34	0.57	0.84	1.15	1.50	1.90	2.33	2.79	3.30	3.84	4.42	5.02	5.66	6.33	7.03	7.76	8.52	9.30	10.10
7.00	0.48	0.80	1.17	1.61	2.09	2.64	3.23	3.87	4.57	5.31	6.09	6.92	7.78	8.69	9.63	10.60	11.61	12.64	13.70
8.00	0.64	1.07	1.57	2.15	2.80	3.52	4.31	5.16	6.08	7.05	8.08	9.16	10.28	11.45	12.66	13.90	15.18	16.49	17.82
9.00	0.84	1.40	2.05	2.80	3.65	4.58	5.59	6.69	7.86	9.10	10.41	11.77	13.19	14.65	16.16	17.70	19.28	20.88	22.50
10.00	1.08	1.79	2.62	3.57	4.64	5.82	7.11	8.49	9.96	11.51	13.13	14.82	16.57	18.36	20.20	22.07	23.97	25.89	27.83
11.00	1.36	2.25	3.29	4.49	5.83	7.30	8.90	10.61	12.42	14.34	16.33	18.39	20.52	22.70	24.92	27.17	29.44	31.73	34.03
12.00	1.70	2.80	4.09	5.57	7.23	9.05	11.03	13.14	15.37	17.72	20.16	22.68	25.27	27.92	30.62	33.36	36.13	38.94	44.72

Note: Table shows aggregate R&D subsidies as a share of total R&D spending in the model with separated product markets (definition 2). Descriptions of the models can be found in Section 3.2. The parameter $(\lambda-1)*100$ (rows) measures the increase in quality in each innovative step in percent and the parameter $w*100$ (columns) measures the external benefit of research that accrues to the government in percent of the private benefit from innovation.

Table 22: Simulation Results: R&D subsidies as a share of R&D spending under integrated product markets

$(\lambda-1)*100$ \ $w*100$	0.00	10.00	20.00	30.00	40.00	50.00	60.00	70.00	80.00	90.00	100.00	110.00	120.00	130.00	140.00	150.00	160.00	170.00	180.00
1.00	0.18	3.26	6.18	8.94	11.52	13.95	16.23	18.38	20.39	22.29	24.08	25.77	27.37	28.88	30.31	31.67	32.96	34.19	35.37
2.00	0.41	3.52	6.48	9.26	11.87	14.33	16.63	18.79	20.82	22.74	24.54	26.25	27.86	29.38	30.83	32.20	33.50	34.74	35.93
3.00	0.67	3.84	6.83	9.65	12.30	14.79	17.12	19.31	21.37	23.31	25.14	26.86	28.49	30.04	31.50	32.90	34.22	35.78	38.08
4.00	0.97	4.20	7.25	10.12	12.81	15.33	17.71	19.93	22.03	24.00	25.86	27.62	29.28	30.86	32.35	34.57	37.17	39.58	41.83
5.00	1.32	4.61	7.72	10.65	13.40	15.97	18.39	20.67	22.81	24.82	26.72	28.52	30.22	32.57	35.54	38.29	40.82	43.17	45.36
6.00	1.71	5.08	8.26	11.25	14.07	16.70	19.18	21.51	23.71	25.77	27.72	29.61	33.07	36.23	39.14	41.82	44.29	46.58	48.72
7.00	2.14	5.60	8.86	11.94	14.82	17.53	20.08	22.48	24.73	26.86	29.54	33.26	36.65	39.75	42.59	45.20	47.62	49.86	51.94
8.00	2.61	6.17	9.53	12.70	15.67	18.47	21.09	23.57	25.90	29.13	33.16	36.80	40.12	43.16	45.94	48.50	50.86	53.06	55.09
9.00	3.13	6.80	10.27	13.54	16.62	19.51	22.23	24.80	28.35	32.74	36.69	40.27	43.53	46.51	49.24	51.76	54.09	56.24	58.25
10.00	3.70	7.50	11.10	14.49	17.68	20.69	23.52	27.16	31.99	36.31	40.20	43.73	46.94	49.89	52.59	55.08	57.39	59.54	61.55
11.00	4.32	8.27	12.01	15.54	18.87	22.01	25.51	30.86	35.62	39.90	43.76	47.26	50.47	53.41	56.13	58.66	61.03	63.27	65.43
12.00	4.99	9.11	13.02	16.72	20.22	23.52	29.31	34.62	39.36	43.63	47.52	51.07	54.36	57.45	60.43	63.57	65.49	67.52	69.34

Note: Table shows aggregate R&D subsidies as a share of total R&D spending in the model with integrated product markets (definition 2). Descriptions of the models can be found in Section 3.2. The parameter $(\lambda-1)*100$ (rows) measures the increase in quality in each innovative step in percent and the parameter $w*100$ (columns) measures the external benefit of research that accrues to the government in percent of the private benefit from innovation.

Table 23: Research subsidy rates for technology follower firms under separated product markets

$(\lambda-1)*100$ \ $w*100$	0.00	10.00	20.00	30.00	40.00	50.00	60.00	70.00	80.00	90.00	100.00	110.00	120.00	130.00	140.00	150.00	160.00	170.00	180.00
1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	9.51

Note: Table shows research subsidy rates offered to technology follower firms in the model with separated product markets (see definition 1). The parameter $(\lambda-1)*100$ (rows) measures the increase in quality in each innovative step in percent and the parameter $w*100$ (columns) measures the external benefit of research that accrues to the government in percent of the private benefit from innovation.

Table 24: Research subsidy rates for technology leader firms under separated product markets

$(\lambda-1)*100$ \ $w*100$	0.00	10.00	20.00	30.00	40.00	50.00	60.00	70.00	80.00	90.00	100.00	110.00	120.00	130.00	140.00	150.00	160.00	170.00	180.00
1.00	23.67	32.08	38.83	44.35	48.96	52.87	56.22	59.12	61.67	63.91	65.91	67.70	69.31	70.76	72.09	73.30	74.41	75.43	76.37
2.00	24.11	32.52	39.25	44.76	49.35	53.24	56.58	59.47	62.00	64.23	66.22	67.99	69.59	71.04	72.35	73.56	74.66	75.67	76.61
3.00	24.57	32.97	39.69	45.18	49.76	53.64	56.95	59.83	62.35	64.57	66.54	68.31	69.89	71.33	72.64	73.83	74.92	75.93	76.86
4.00	25.04	33.44	40.14	45.63	50.19	54.05	57.35	60.21	62.72	64.92	66.88	68.64	70.22	71.64	72.94	74.12	75.21	76.21	77.13
5.00	25.53	33.93	40.62	46.09	50.64	54.48	57.77	60.61	63.10	65.30	67.25	68.99	70.56	71.97	73.26	74.44	75.51	76.50	77.42
6.00	26.04	34.43	41.12	46.58	51.11	54.93	58.21	61.04	63.52	65.70	67.63	69.36	70.92	72.33	73.61	74.77	75.84	76.83	77.73
7.00	26.57	34.96	41.64	47.08	51.60	55.41	58.67	61.49	63.95	66.12	68.05	69.77	71.31	72.71	73.98	75.14	76.20	77.17	78.07
8.00	27.12	35.52	42.19	47.62	52.13	55.92	59.17	61.97	64.42	66.58	68.49	70.20	71.74	73.12	74.38	75.53	76.59	77.56	78.45
9.00	27.69	36.10	42.77	48.19	52.68	56.46	59.70	62.49	64.92	67.07	68.97	70.67	72.20	73.58	74.83	75.97	77.02	77.98	78.86
10.00	28.29	36.71	43.37	48.79	53.27	57.04	60.26	63.05	65.47	67.60	69.50	71.18	72.70	74.07	75.32	76.46	77.50	78.45	79.33
11.00	28.92	37.35	44.02	49.43	53.91	57.67	60.88	63.65	66.07	68.19	70.08	71.76	73.27	74.64	75.88	77.01	78.05	79.00	79.88
12.00	29.57	38.03	44.71	50.12	54.59	58.35	61.56	64.32	66.73	68.85	70.73	72.41	73.92	75.29	76.53	77.67	78.71	79.67	80.77

Note: Table shows research subsidy rates offered to technology follower firms in the model with separated product markets (see definition 1). The parameter $(\lambda-1)*100$ (rows) measures the increase in quality in each innovative step in percent and the parameter $w*100$ (columns) measures the external benefit of research that accrues to the government in percent of the private benefit from innovation.

Table 25: Research subsidy rates for technology follower firms under integrated product markets in the “leveled” state of competition.

$(\lambda-1)*100$ \ $w*100$	0.00	10.00	20.00	30.00	40.00	50.00	60.00	70.00	80.00	90.00	100.00	110.00	120.00	130.00	140.00	150.00	160.00	170.00	180.00
1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.95	4.55
4.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.49	6.45	10.11	13.52
5.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.19	6.60	10.67	14.43	17.92	21.15
6.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	5.12	9.71	13.93	17.81	21.40	24.72	27.81
7.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.89	7.19	12.02	16.42	20.46	24.18	27.61	30.79	33.74
8.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.88	8.53	13.63	18.27	22.52	26.40	29.98	33.28	36.33	39.17
9.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.14	9.20	14.65	19.59	24.09	28.19	31.95	35.41	38.61	41.57	44.32
10.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.67	9.25	15.13	20.44	25.25	29.62	33.63	37.30	40.69	43.83	46.75	49.48
11.00	0.00	0.00	0.00	0.00	0.00	0.00	1.46	8.68	15.10	20.87	26.07	30.80	35.12	39.09	42.77	46.19	49.41	52.48	55.45
12.00	0.00	0.00	0.00	0.00	0.00	0.00	7.46	14.57	20.91	26.63	31.84	36.62	41.06	45.26	49.39	53.97	71.14	73.34	75.32

Note: Table shows research subsidy rates offered to technology follower firms in the model with integrated product markets in the “leveled” state of competition (see definition 3), meaning that both countries host exactly one technology leader firm and one technology follower firm. The parameter $(\lambda-1)*100$ (rows) measures the increase in quality in each innovative step in percent and the parameter $w*100$ (columns) measures the external benefit of research that accrues to the government in percent of the private benefit from innovation.

Table 26: Research subsidy rates for technology leader firms under integrated product markets in the "leveled" state of competition.

$w * 100$ / $(\lambda - 1) * 100$	0.00	10.00	20.00	30.00	40.00	50.00	60.00	70.00	80.00	90.00	100.00	110.00	120.00	130.00	140.00	150.00	160.00	170.00	180.00
1.00	0.49	8.91	15.79	21.53	26.41	30.59	34.24	37.44	40.28	42.82	45.10	47.17	49.05	50.77	52.35	53.80	55.15	56.40	57.57
2.00	0.98	9.40	16.27	22.00	26.85	31.03	34.66	37.85	40.68	43.21	45.49	47.54	49.42	51.13	52.70	54.15	55.48	56.73	57.89
3.00	1.49	9.89	16.75	22.47	27.32	31.48	35.10	38.28	41.10	43.61	45.88	47.92	49.79	51.49	53.05	54.49	55.82	57.15	58.66
4.00	2.00	10.39	17.25	22.95	27.78	31.94	35.54	38.71	41.52	44.02	46.28	48.32	50.17	51.86	53.41	55.10	56.84	58.45	59.94
5.00	2.52	10.91	17.76	23.45	28.27	32.41	36.00	39.16	41.95	44.45	46.69	48.72	50.56	52.48	54.50	56.36	58.08	59.67	61.16
6.00	3.07	11.45	18.29	23.97	28.77	32.90	36.48	39.63	42.41	44.89	47.13	49.15	51.54	53.72	55.73	57.57	59.28	60.87	62.34
7.00	3.63	12.01	18.83	24.51	29.30	33.42	36.99	40.12	42.89	45.36	47.79	50.40	52.77	54.94	56.94	58.78	60.47	62.05	63.51
8.00	4.21	12.60	19.41	25.07	29.86	33.96	37.52	40.64	43.40	46.19	49.05	51.65	54.01	56.18	58.16	59.99	61.68	63.25	64.71
9.00	4.83	13.20	20.01	25.67	30.45	34.54	38.09	41.20	44.31	47.48	50.34	52.93	55.29	57.44	59.43	61.26	62.95	64.52	65.98
10.00	5.46	13.85	20.66	26.31	31.08	35.17	38.71	42.13	45.65	48.82	51.67	54.27	56.63	58.80	60.79	62.63	64.33	65.92	67.40
11.00	6.14	14.54	21.35	27.01	31.78	35.86	39.57	43.53	47.07	50.24	53.12	55.72	58.11	60.31	62.34	64.22	65.99	67.65	69.24
12.00	6.86	15.28	22.11	27.77	32.55	36.63	41.06	45.05	48.61	51.83	54.75	57.42	59.89	62.20	64.41	66.66	65.18	67.50	69.59

Note: Table shows research subsidy rates offered to technology leader firms in the model with integrated product markets in the "leveled" state of competition (see definition 3), meaning that both countries host exactly one technology leader firm and one technology follower firm. The parameter $(\lambda - 1) * 100$ (rows) measures the increase in quality in each innovative step in percent and the parameter $w * 100$ (columns) measures the external benefit of research that accrues to the government in percent of the private benefit from innovation.

Table 27: Research subsidy rates for technology follower firms under integrated product markets in the "unleveled" state of competition.

$w * 100$ / $(\lambda - 1) * 100$	0.00	10.00	20.00	30.00	40.00	50.00	60.00	70.00	80.00	90.00	100.00	110.00	120.00	130.00	140.00	150.00	160.00	170.00	180.00
1.00	23.64	31.66	38.15	43.52	48.02	51.87	55.18	58.07	60.61	62.85	64.86	66.66	68.29	69.76	71.11	72.34	73.47	74.51	75.47
2.00	24.06	32.06	38.54	43.89	48.39	52.22	55.52	58.40	60.93	63.16	65.16	66.95	68.57	70.04	71.38	72.60	73.72	74.76	75.71
3.00	24.48	32.48	38.95	44.29	48.77	52.59	55.88	58.74	61.26	63.49	65.48	67.26	68.87	70.33	71.66	72.87	73.99	75.04	76.06
4.00	24.92	32.91	39.37	44.70	49.17	52.98	56.26	59.11	61.62	63.83	65.81	67.59	69.19	70.64	71.96	73.22	74.41	75.51	76.52
5.00	25.37	33.36	39.81	45.13	49.59	53.38	56.65	59.49	61.99	64.20	66.17	67.93	69.53	71.01	72.41	73.70	74.87	75.95	76.95
6.00	25.84	33.82	40.26	45.57	50.02	53.81	57.06	59.90	62.38	64.58	66.54	68.30	69.98	71.50	72.89	74.16	75.32	76.38	77.37
7.00	26.32	34.30	40.74	46.04	50.48	54.26	57.50	60.33	62.80	65.00	66.98	68.82	70.48	71.98	73.35	74.60	75.74	76.80	77.77
8.00	26.81	34.80	41.23	46.53	50.96	54.73	57.97	60.78	63.25	65.48	67.51	69.32	70.96	72.44	73.79	75.03	76.15	77.19	78.15
9.00	27.33	35.31	41.75	47.04	51.47	55.23	58.46	61.27	63.77	66.02	68.02	69.81	71.43	72.89	74.22	75.43	76.54	77.55	78.49
10.00	27.86	35.85	42.29	47.58	52.01	55.77	58.99	61.82	64.32	66.54	68.51	70.28	71.87	73.30	74.60	75.78	76.86	77.84	78.74
11.00	28.40	36.42	42.86	48.16	52.58	56.34	59.57	62.38	64.85	67.03	68.97	70.70	72.24	73.63	74.88	75.99	76.99	77.86	78.59
12.00	28.96	37.00	43.46	48.76	53.20	56.96	60.15	62.91	65.33	67.46	69.33	70.97	72.41	73.64	74.60	74.99	64.07	66.01	67.76

Note: Table shows research subsidy rates offered to technology follower firms in the model with integrated product markets in the "unleveled" state of competition (see definition 3), meaning that one country hosts both technology leader firms and one country hosts both technology follower firms. The parameter $(\lambda - 1) * 100$ (rows) measures the increase in quality in each innovative step in percent and the parameter $w * 100$ (columns) measures the external benefit of research that accrues to the government in percent of the private benefit from innovation.

Table 28: Research subsidy rates for technology leader firms under integrated product markets in the "unleveled" state of competition.

$(\lambda - 1) * 100$ \ $w * 100$	0.00	10.00	20.00	30.00	40.00	50.00	60.00	70.00	80.00	90.00	100.00	110.00	120.00	130.00	140.00	150.00	160.00	170.00	180.00
1.00	49.99	54.87	58.90	62.29	65.17	67.66	69.82	71.71	73.39	74.88	76.21	77.42	78.51	79.50	80.41	81.24	82.00	82.71	83.36
2.00	49.97	54.87	58.92	62.31	65.20	67.69	69.85	71.75	73.43	74.92	76.26	77.46	78.55	79.55	80.45	81.28	82.05	82.75	83.40
3.00	49.95	54.86	58.92	62.33	65.23	67.72	69.89	71.79	73.47	74.96	76.30	77.51	78.60	79.59	80.50	81.33	82.09	82.79	83.43
4.00	49.92	54.85	58.92	62.34	65.25	67.75	69.92	71.82	73.51	75.00	76.34	77.55	78.64	79.64	80.54	81.36	82.10	82.79	83.43
5.00	49.87	54.83	58.92	62.35	65.26	67.77	69.95	71.86	73.54	75.04	76.38	77.59	78.69	79.66	80.54	81.35	82.09	82.77	83.41
6.00	49.82	54.80	58.91	62.35	65.27	67.79	69.97	71.89	73.58	75.08	76.42	77.63	78.68	79.64	80.51	81.31	82.05	82.74	83.37
7.00	49.76	54.76	58.89	62.35	65.28	67.80	70.00	71.91	73.61	75.11	76.44	77.59	78.64	79.59	80.46	81.26	81.99	82.67	83.30
8.00	49.68	54.71	58.86	62.33	65.28	67.81	70.01	71.94	73.64	75.10	76.38	77.52	78.56	79.51	80.37	81.17	81.90	82.57	83.19
9.00	49.59	54.65	58.82	62.31	65.27	67.82	70.02	71.96	73.60	75.01	76.27	77.41	78.44	79.38	80.24	81.02	81.74	82.40	83.01
10.00	49.47	54.57	58.76	62.28	65.25	67.81	70.03	71.91	73.47	74.87	76.12	77.24	78.26	79.17	80.01	80.77	81.46	82.09	82.66
11.00	49.34	54.47	58.69	62.22	65.22	67.79	69.98	71.73	73.27	74.65	75.87	76.95	77.93	78.80	79.58	80.26	80.85	81.35	81.71
12.00	49.17	54.34	58.59	62.15	65.16	67.75	69.73	71.45	72.95	74.26	75.41	76.40	77.24	77.90	78.30	78.02	62.23	64.23	66.03

Note: Table shows research subsidy rates offered to technology leader firms in the model with integrated product markets in the "unleveled" state of competition (see definition 3), meaning that one country hosts both technology leader firms and one country hosts both technology follower firms. The parameter $(\lambda - 1) * 100$ (rows) measures the increase in quality in each innovative step in percent and the parameter $w * 100$ (columns) measures the external benefit of research that accrues to the government in percent of the private benefit from innovation.

References

Chapter 1

[1] Acemoglu, D., Akcigit, U., Hanley, D. & Kerr, W., 2014. Transition to Clean Technology. Working Paper.

[2] Akcigit, U., Hanley, D. & Serrano-Verde, N., 2013. Back to Basics: Basic Research Spillovers, Innovation Policy and Growth. Working paper.

[3] Almus, M. & Czarnitzki, D., 2003. The effects of public R&D subsidies on firm's innovation activities. The case of Eastern Germany. Journal of Business & Economic Statistics, 21(2), 226-236.

[4] Arnold, E., Falk, M., Falk, R., Knoll, N., Leo, H., & Schwarz, G., 2004. Evaluation of the Austrian Industrial Research Promotion Fund (FFF) and the Austrian Science Fund (FWF). Synthesis Report.
URL: www.ffg.at/sites/default/files/downloads/page/bpevaluierung03synthesisrep.pdf

[5] Becker, B., 2015. Public R&D policies and private R&D investment: a survey of the empirical evidence. Journal of Economic Surveys, 29(5), 917-942.

[6] Bloom, N., Schankerman, M., & Van Reenen, J., 2013. Identifying Technology Spillovers and Product Market Rivalry. Econometrica, 81(4), 1347-1393.

[7] Bérubé, C. & Mohnen, P., 2009. Are firms that receive R&D subsidies more innovative? Canadian Journal of Economics, 42(1), 206-225.

[8] Berkes, E. & Gaetani, R., 2016. The Geography of Unconventional Innovation. Working Paper.

[9] Besanko, D. & Wu, J., 2010. Government Subsidies for Research Programs Facing If and When Uncertainty. Working Paper.

[10] Blundell, R. & Griffith, R., & Van Reenen, J., 1999. Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms. Review of Economic Studies, 66(3), 529-554.

[11] Branstetter, L. & Sakakibara, M., 2002. When Do Research Consortia Work Well and Why? Evidence from Japanese Panel Data. The American Economic Review, 92(1), 143-159.

[12] Bronzini, R. & Iachini, E., 2014. Are Incentives for R&D Effective? Evidence from a Regression Discontinuity Approach. American Economic Journal: Economic Policy, 6(4), 100-134.

[13] Bronzini, R. & Piselli, P., 2016. The impact of R&D subsidies on firm innovation. Research Policy, 45(2), 442-457.

[14] Bryan, K. & Lemus, J., 2016. The Direction of Innovation. Working paper.

[15] Cappelen, A., Raknerud, A. & Rybalka, M., 2012. The effect of R&D tax credits on-patenting and innovations. Research Policy, 41(2), 334-345.

[16] Czarnitzki, D., Hanel P. & Rosa, JM, 2012. Evaluating the impact of R&D tax credits on innovation: A microeconometric study on Canadian firms. Research Policy, 40(2), 217-229.

[17] Dechezleprêtre, A., Einiö, E., Martin, R., Nguyen, K. T. & Van Reenen, J., 2016. Do tax incentives for research increase firm innovation? An RD design for R&D. Working paper.

[18] Einiö, E., 2014. R&D subsidies and company performance: evidence from geographic variation in government funding based on the ERDF population-density rule. Review of Economics and Statistics, 96(4), 710-728.

[19] González, X., Jaumandreu, J. & Pazó, C., 2005. Barriers to innovation and subsidy effectiveness. RAND Journal of Economics, 36(4), 930-950.

[20] Hall, B. H., 2002. The financing of research and development. Oxford Review of Economic Policy, 18(1), 35-51.

[21] Hall, B. H., Jaffe, A. B. & Trajtenberg, M., 2005. Market Value and Patent Citations. RAND Journal of Economics, 36(1), 16-38.

[22] Howell, S., 2017. Financing Innovation: Evidence from R&D Grants. American Economic Review, 107(4), 1136164.

[23] Hopenhayn, H. & Squintani, F., 2016. The Direction of Innovation. Working paper.

[24] Hussinger, K., 2008. R&D and subsidies at the firm level: an application of parametric and semiparametric two-step selection models. Journal of Applied Econometrics, 23(6), 729-747.

[25] Jaffe, A. B., 1986. Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value. American Economic Review, 76(5), 984-1001.

[26] Jones, C. & William, J., 1998. Measuring the Social Rate of Return to R&D. Quarterly Journal of Economics, 113(4), 1119-1135.

[27] Jovanovic, B. & Rob, R., 1990. Long waves and short waves: Growth through intensive and extensive search. Econometrica, 58(6), 1391-1409.

[28] Lach, S., 2002. Do R&D subsidies stimulate or displace private R&D? Evidence from Israel. Journal of Industrial Economics, 50(4), 369-390.

[29] Lerner, J., 2000. The government as venture capitalist: the long-run impact of the SBIR program. The Journal of Private Equity, 3(2), 55-78.

[30] Lerner, J., 2009. Boulevard of broken dreams: why public efforts to boost entrepreneurship and venture capital have failed-and what to do about it. Princeton University Press.

[31] Lichtenberg, F. R., 1988. The private R&D investment response to federal design and technical competitions. American Economic Review, 78(1), 550-559.

[32] Nelson, R. R., 1959. The Simple Economics of Basic Scientific Research. Journal of Political Economy 67(3), 297-306.

[33] OECD 2014. Science, Technology and Industry Outlook.

[34] Takalo, T., Tanayama, T. & Toivanen O., 2013. Estimating the benefits of targeted R&D subsidies. The Review of Economics and Statistics, 95(1), 255-272.

[35] Trajtenberg, M., 1990. A Penny for Your Quotes: Patent Citations and the Value of Innovations. RAND Journal of Economics, 21(1), 172-187.

[36] Wallsten, S. J., 2000. The effects of government-industry R&D programs on private R&D: the case of the small business innovation research program. RAND Journal of Economics 31(1): 82-100.

[37] Weitzman, M., 1979. Optimal Search for the Best Alternative. Econometrica, 47(3), 641-654.

[38] Zúñiga-Vicente, J., Alonso-Borrego, C., Forcadell, F.J., Galàn, J.I., 2014. Assessing the effect of public subsidies on firm R&D investment: a survey. Journal of Economic Surveys 28(1), 36-67.

Chapter 2

[39] Akcigit, U., Baslandze, S., and Stantcheva, S., 2016. Taxation and the International Mobility of Inventors. American Economic Review, 106(10), 2930-2981.

[40] Albouy, D., 2008. Are Big Cities Bad Places to Live? Estimating of Life across Metropolitan Areas. NBER working paper 14472.

[41] Albouy, D., Leibovici, F. & Warman, C., 2013. Quality of Life, Firm Productivity, and the Value of Amenities across Canadian Cities. Canadian Journal of Economics, 49(2), 379-411.

[42] Albouy, D., 2016. What are Cities Worth? Land Rents, Local Productivity, and the Total Value of Amenities. Review of Economics and Statistics, 98(3), 477-487.

[43] Bayer, P., Ferreira, F. & McMillan, R., 2007. A Unified Frame for Measuring Preferences for Schools and Neighboorhoods. Journal of Political Economy, 115(4), 588-638.

[44] Bakija, J., & Slemrod, J., 2004. Do the Rich Flee from High State Taxes? Evidence from Federal Estate Tax Returns. NBER working paper 10645.

[45] Bereitschaft, B. & Cammack, R., 2015. Neighborhood diversity and the creative class in Chicago. Applied Geography, 63, 166-183.

[46] Black, S., 1999. Do better schools matter? Parental valuation of elementary education. Quarterly Journal of Economics, 114(2), 577-599.

[47] Brown, M. W. & Scott, D. M., 2012. Human Capital Location Choice: Accounting for Amenities and Thick Labor Markets. Journal of Regional Science, 52(5), 787-808.

[48] Cohen, R., Lai, A. & Steindel, C., 2011. The Effects of Marginal Tax Rates on Interstate Migration in the U.S. Working Paper at New Jersey Department of the Treasury.

[49] Cullen, J. B., Levitt, S. D., 1999. Crime, Urban Flight, and the Consequences for Cities. Review of Economics and Statistics, 81(2), 159-169.

[50] Dahl, M. S. & Sorensen, O., 2010. The Migration of Technical Workers. Journal of Urban Economics, 67(1), 33-45.

[51] Dorfman, J. H., Partridge, M. D. & Galloway, H., 2011. Do Natural Amenities Attract High-Tech Jobs? Evidence from a Smoothed Bayesian Spatial Model. Spatial Economic Analysis, 6(4), 397-422.

[52] Dorner, M., Harhoff, D., Hinz, T., Hoisl, K. & Bender, S., 2017. Social Ties for Labor Market Access: Lessons from the Migration of East German Inventors. CEPR Discussion Paper DP11601.

[53] Faggian, A., & McCann, P., 2009. Human Capital, Graduate Migration and Innovation in British Regions. Cambridge Journal of Economics, 33(2), 317-333.

[54] Florida, R., 2002. The Rise of the Creative Class: And how it's transforming work, leisure, community and everyday life. New York, Perseus Book Group.

[55] Frenkel, A., Bendit, E. & Kaplan, S., 2013. Residential location choice of knowledge-workers: The role of amenities, workplace and lifestyle. Cities, 35, 33-41.

[56] Glaeser, E. L., 2005. Smart Growth: Education, Skilled Workers and the Future of Cold-Weather Cities. Cambridge, MA: Harvard University, Kennedy School, Policy Brief PB-2005-1.

[57] Glaeser, E. L. & Tobio, K., 2008. The Rise of the Sunbelt. Southern Economic Journal, 74(3), 610-643.

[58] Gottlieb, P. D. & Joseph, G., 2006. College-to-Work Migration of Technology Graduates and Holders of Doctorates within the United States. Journal of Regional Science, 46(4), 627-659.

[59] Kleven, H. J., Landais, C. & Saez, E., 2013. Taxation and International Migration of Superstars: Evidence from the European Football Market. American Economic Review, 103(5), 1892-1924.

[60] Kirchgässner, G. & Pommerehne, W.W., 2006. Tax harmonization and tax competition in the European Union: Lessons from Switzerland. Journal of Public Economics, 60(3), 351-371.

[61] Lawton, P., Murphy, E. & Redmond, D., 2013. Residential preferences of the 'creative class'? Cities, 31, 47-56.

[62] Liebig, T., Puhani, P.A. & Sousa-Poza, A., 2007. Taxation and Internal Migration: Evidence from the Swiss Census Using Community-Level Variation in Income Tax Rates. Journal of Regional Science, 47(4), 807-836.

[63] McFadden, D., 1978. Modelling the Choice of Residential Location. In: Karlqvist, A., Lundqvist, L., Snickars, F., Weibull, J. (ed). Spatial Interaction Theory and Planning Models, Amsterdam, 75-96.

[64] Miguélez, E. & Fink, C., 2013. Measuring the International Mobility of Inventors: A New Database. World Intellectual Property Organization Research Working Paper No. 8

[65] Miguélez, E. & Moreno, R., 2013. What attracts Knowledge Workers? The Role of Space and Social Networks. Journal of Regional Science, 54(10), 33-60.

[66] Moretti, E. & Wilson, D. J., 2017. The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists. American Economic Review, 107(10), 1858-1903.

[67] Schmidheiny, K., 2006. Income Segregation and Local Progressive Taxation: Empirical Evidence from Switzerland. Journal of Public Economics, 90 (3), 429-58.

[68] Scott, A. J., 2010. Jobs or Amenities? Destination Choices of Migrant Engineers in the USA. Papers in Regional Science, 89(1), 43-63.

[69] Wojan, T. R., Lambert, D. M. & McGranahan, D. A., 2007. Emoting with their feet: Bohemian attraction to creative milieu. Journal of Economic Geography, 7, 711-736.

[70] Young, C., & Varner, C., 2011. Millionaire Migration and State Taxation of Top Incomes: Evidence from a Natural Experiment. National Tax Journal, 64(2), 255-283.

Chapter 3

[71] Aghion, P. & Howitt, P. 1992. A Model of Growth through Creative Destruction. Econometrica, 60(2), p. 323-351

[72] Garcia Pires, A. J. 2015. Competitiveness-shifting effects and the prisoner's dilemma in international R&D subsidy wars. International Economics, 142, p. 32-49

[73] Haaland, J. I. & Kind, H. J. 2008. R&D policies, trade and process innovation. Journal of International Economics, 74, p. 170-187

[74] Hammadou, H., Paty S. & Savona, M. 2014. Strategic interactions in public R&D across European countries: A spatial econometric analysis. Research Policy, 43, p. 1217-1226

[75] Impullitti, G. 2010. International Competition and U.S. R&D Subsidies: A Quantitative Welfare Analysis. International Economic Review, 51 (4), p. 1127-1158

[76] Kondo, H. 2013. International R&D subsidy competition, industrial agglomeration and growth. Journal of International Economics, 89, p. 233-251

[77] OECD Science. Technology and Industry Outlook 2014. OECD

[78] Warda, J. 2001. Measuring the value of RD tax treatment in OECD countries. STI Review: Special issue on new science and technology indicators, 27, OECD

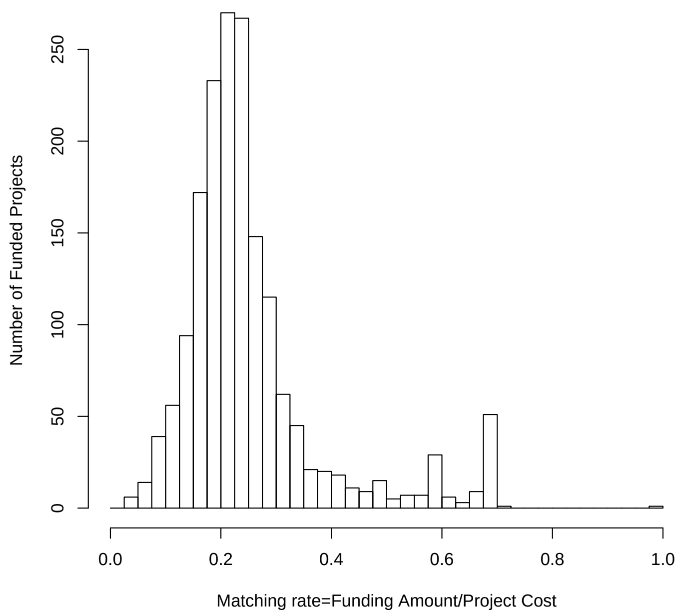
5 Appendix

Chapter 1

Table 29: Number of funding applications per firm

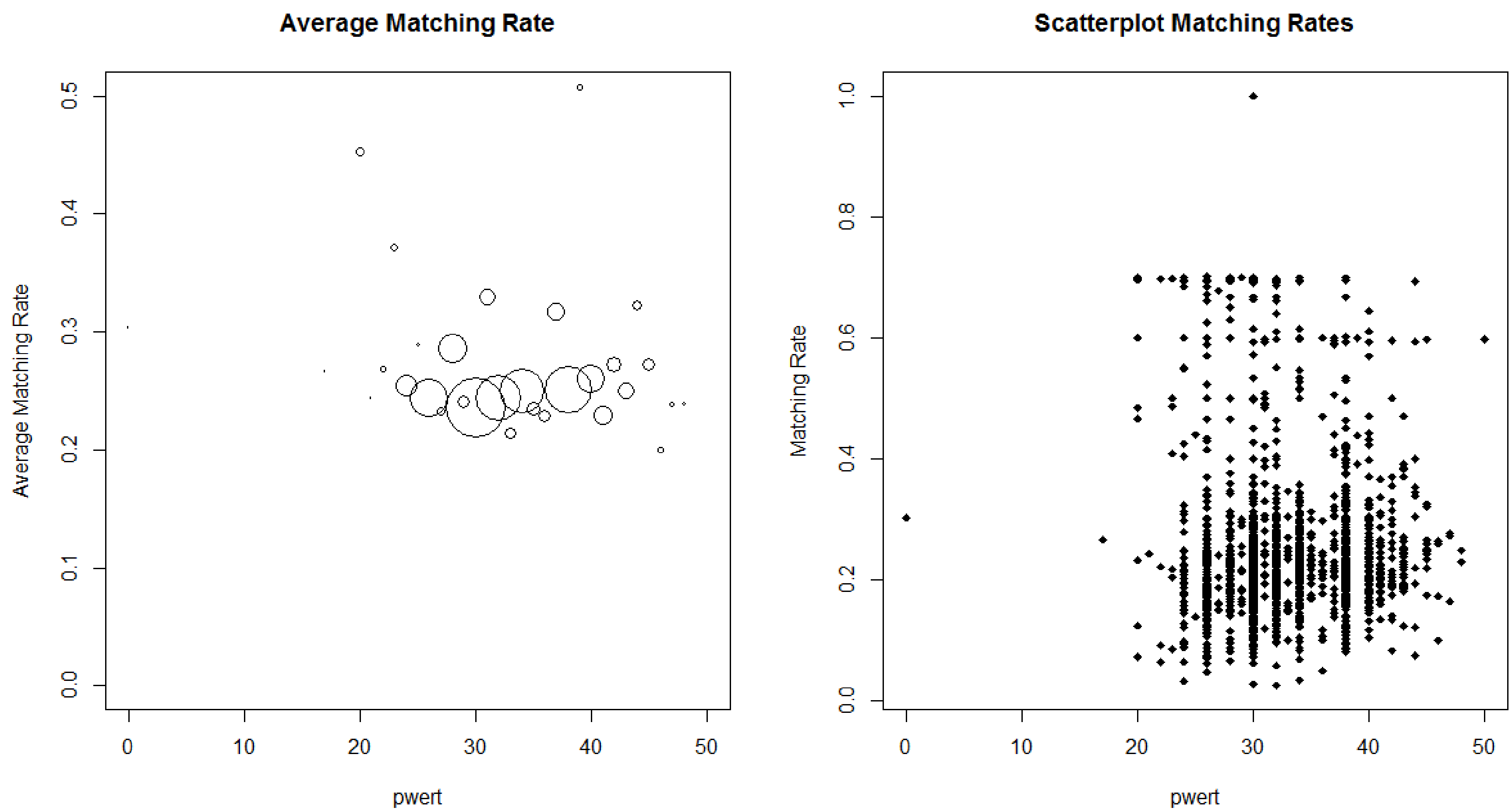
# of applications by the same firm in the given period	2002	2003	2004	2005	2002-2005
1	763	761	707	691	1419
2	108	137	97	129	424
3	30	26	26	23	200
4	13	6	8	9	109
5	10	4	5	6	54
6	0	4	3	3	38
7	2	1	2	1	20
8	1	0	1	3	14
9	1	0	0	0	14
>=10	3	4	2	2	30

Figure 16: Distribution of matching rates conditional on funding approval in the baseline sample



Note: Matching rate is defined as funding amount divided by total cost of the project.

Figure 17: Matching rates by *pwert* score conditional on funding approval

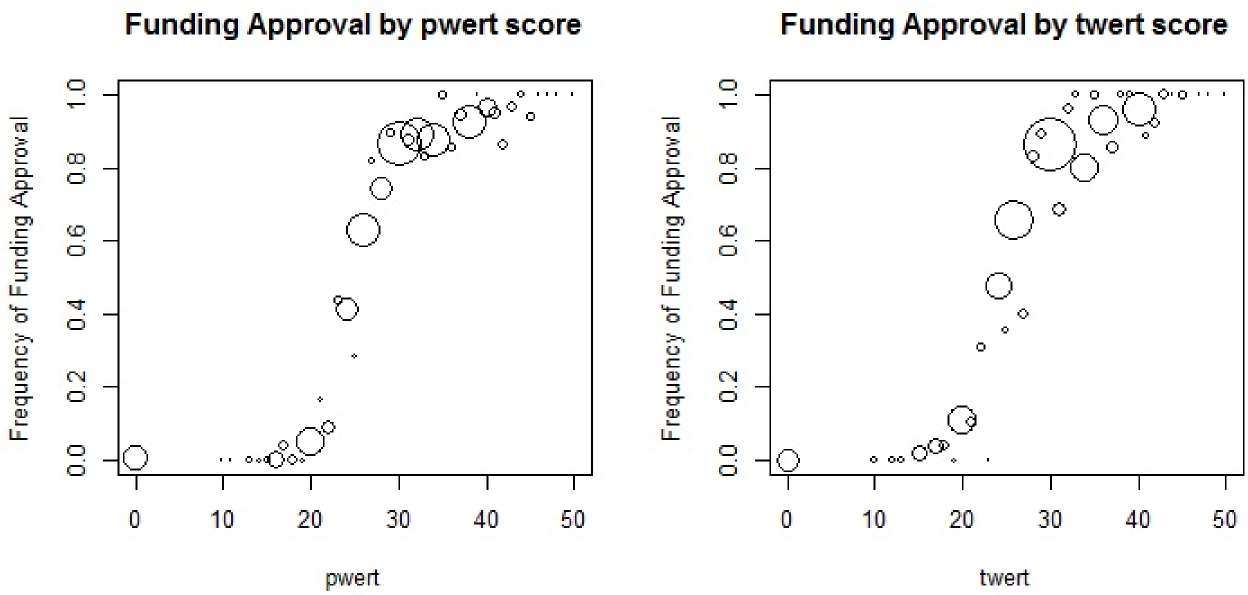


Note: Areas of circles are proportional to number of observations in plot on the left. Matching rate is defined as funding amount divided by total cost of the project.

Table 30: Distribution *pwert* score in the baseline sample

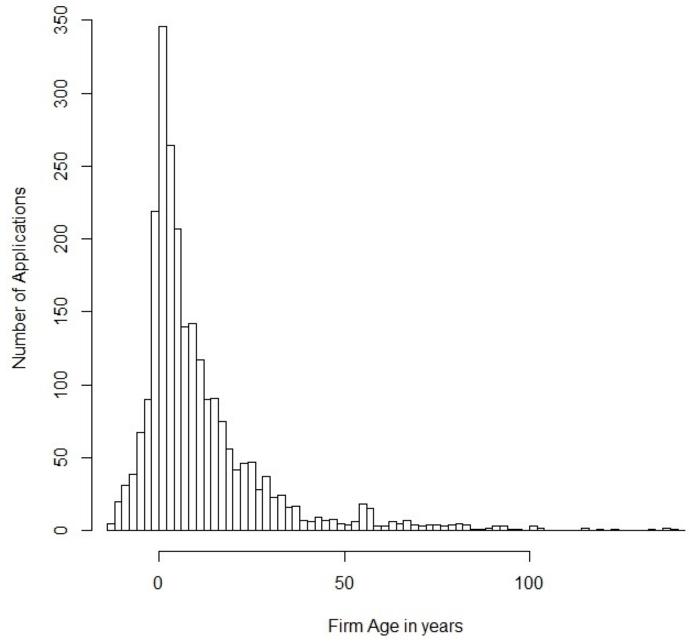
pwert	#obs	pwert	#obs	pwert	#obs	pwert	#obs
0	139	13	12	26	242	39	4
1	0	14	8	27	11	40	89
2	0	15	14	28	124	41	39
3	0	16	70	29	19	42	29
4	0	17	26	30	434	43	29
5	0	18	23	31	33	44	12
6	0	19	5	32	257	45	17
7	0	20	179	33	18	46	4
8	0	21	6	34	242	47	3
9	0	22	46	35	20	48	2
10	1	23	16	36	21	49	0
11	2	24	125	37	38	50	1
12	0	25	7	38	252		

Figure 18: Frequency of funding approval by *pwert* and *twert* score



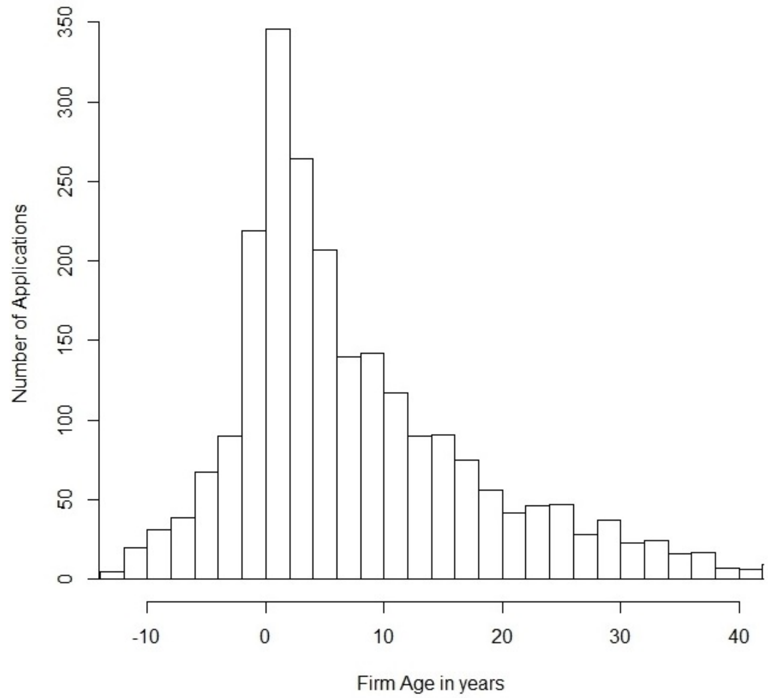
Note: Areas of circles are proportional to number of observations. See section 1.3 for a description of the evaluation scores.

Figure 19: Distribution firm age, all firms



Note: Firm age at the time of the funding application.

Figure 20: Distribution firm age, firms of age less than 40 years



Note: Firm age at the time of the funding application.

Table 31: First stage regression for IV models (1)-(6) in Table 3

Dependent variable: approved, all models estimated with OLS						
	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 1	Model 2	Model 2	Model 3	Model 3
	Cubic Spline	Cubic Spline	2 jumps	2 jumps	3 jumps	3 jumps
pwert	0.00119	0.000725	0.00227***	0.00161	0.00227***	0.00153
	(0.000843)	(0.00167)	(0.000764)	(0.00168)	(0.000764)	(0.00167)
linear spline pwert 21	0.162***	0.160***				
	(0.00644)	(0.00782)				
quad spline pwert 21	-0.00977***	-0.00958***				
	(0.000695)	(0.000748)				
cubic spline pwert 21	0.000192***	0.000187***				
	(2.00e-05)	(2.23e-05)				
linear spline pwert 23			0.00818***	0.00817***	0.00502***	0.00546**
			(0.00186)	(0.00249)	(0.00184)	(0.00258)
jump pwert 23			0.474***	0.474***	0.354***	0.359***
			(0.0296)	(0.0269)	(0.0424)	(0.0337)
jump pwert 26					0.242***	0.236***
					(0.0477)	(0.0317)
jump pwert 27			0.248***	0.246***		
			(0.0319)	(0.0239)		
jump pwert 29					0.169***	0.165***
					(0.0306)	(0.0237)
project cost (in 100K Euro)		-0.00129		-0.00122		-0.00136
		(0.00128)		(0.00128)		(0.00128)
project cost sqr		1.24e-05		1.16e-05		1.30e-05
		(1.16e-05)		(1.17e-05)		(1.16e-05)
EP Patent binary pre		0.0179		0.0215		0.0181
		(0.0200)		(0.0201)		(0.0199)
EP Patents pre		-0.00160		-0.00215		-0.00168
		(0.00254)		(0.00255)		(0.00253)
EP Patents pre sqr		2.04e-05		2.30e-05		1.79e-05
		(2.78e-05)		(2.79e-05)		(2.78e-05)
Nat Patent binary pre		0.00349		0.00394		0.000844
		(0.0187)		(0.0188)		(0.0187)
Nat Patents pre		-2.52e-05		0.000152		0.000182
		(0.00111)		(0.00112)		(0.00111)
Nat Patents pre sqr		7.11e-09		-1.51e-07		-1.80e-07
		(1.04e-06)		(1.05e-06)		(1.04e-06)
Year FE, Sector FE Age FE	NO	YES	NO	YES	NO	YES
Observations	2,619	2,619	2,619	2,619	2,619	2,619
R-squared	0.528	0.543	0.523	0.540	0.530	0.546

Note: Standard errors in parentheses,*** p<0.01, ** p<0.05, * p<0.1

Table 32: The effect of funding approval on the propensity to file a National Patent

Dependent variable: Nat Patent binary post						
	(1) Model 1 IV: Cubic Splines no controls	(2) Model 1 IV: Cubic Splines with controls	(3) Model 2 IV: 2 jumps no controls	(4) Model 2 IV: 2 jumps with controls	(5) Model 3 IV: 3 jumps no controls	(6) Model 3 IV: 3 jumps with controls
approved	0.0428	-0.00106	0.0340	-0.00781	0.0753	0.0154
	(0.0560)	(0.0488)	(0.0515)	(0.0444)	(0.0517)	(0.0448)
Controls	NO	YES	NO	YES	NO	YES
Observations	2,619	2,619	2,619	2,619	2,619	2,619
R-squared	0.032	0.283	0.031	0.283	0.033	0.284

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Mean of dependent variable is 0.25. Same specification (except different dependent variable) as in Table 3.

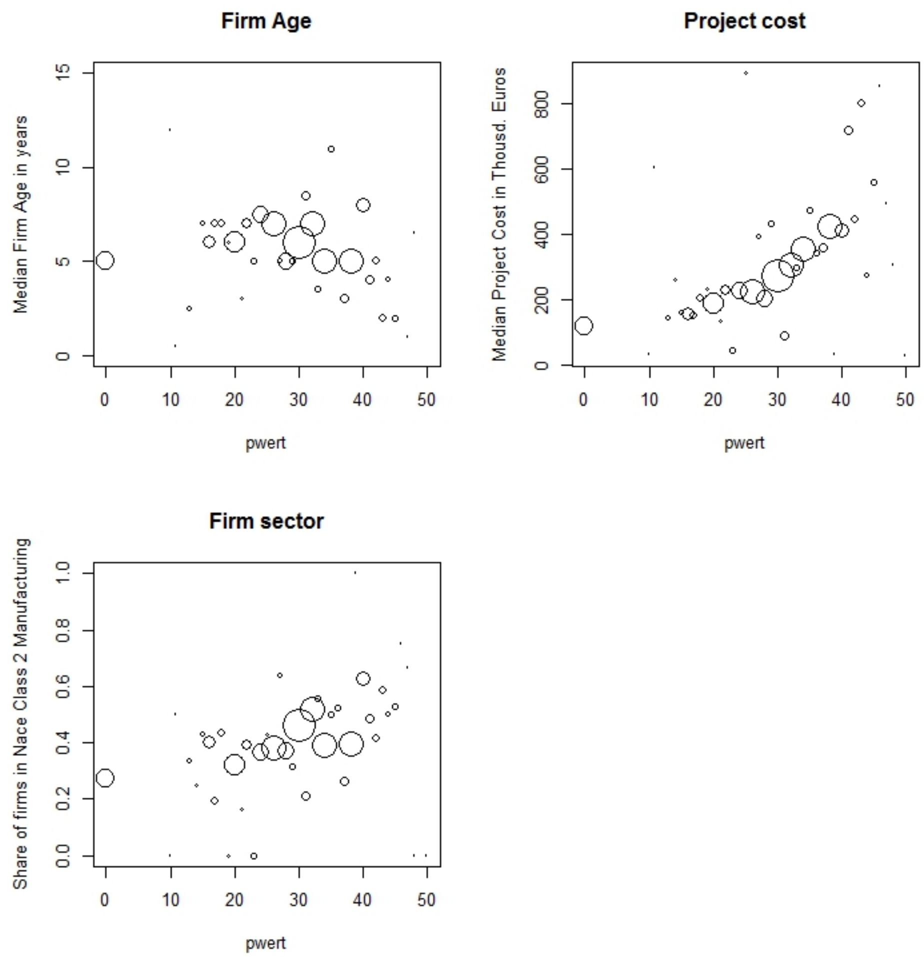
Table 33: The effect of funding approval on the propensity to file a European Patent in the year of the funding application or during the subsequent 4 years (one additional year compared to Table 3)

Dependent variable: EP Patent binary post measured for year of the funding application plus subsequent 4 years						
	(1) Model 1 IV: Cubic Splines	(2) Model 1 IV: Cubic Splines	(3) Model 2 IV: 2 jumps	(4) Model 2 IV: 2 jumps	(5) Model 3 IV: 3 jumps	(6) Model 3 IV: 3 jumps
approved	0.143**	0.120**	0.132**	0.109**	0.179***	0.129***
	(0.0577)	(0.0479)	(0.0534)	(0.0437)	(0.0531)	(0.0435)
Observations	2,619	2,619	2,619	2,619	2,619	2,619
R-squared	0.052	0.348	0.052	0.348	0.050	0.347

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Full set of controls included in all model, but not reported.

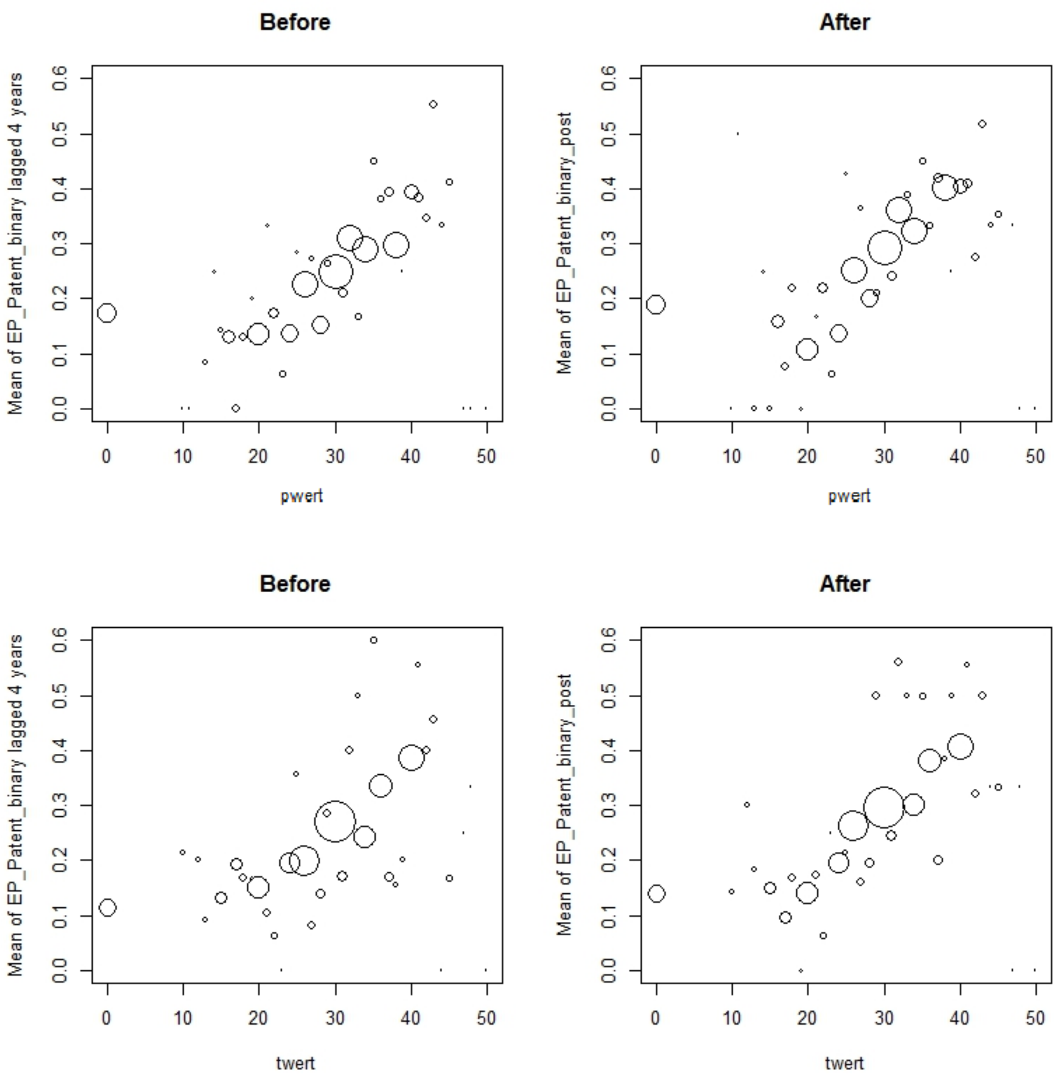
Validity checks for results from section 1.5.1

Figure 22: Median firm age, median project cost and share of firms in Manufacturing of Metal Products, Electronics and Chemical Products (NACE 2 classes 20,22-29) plotted against *pwert* score



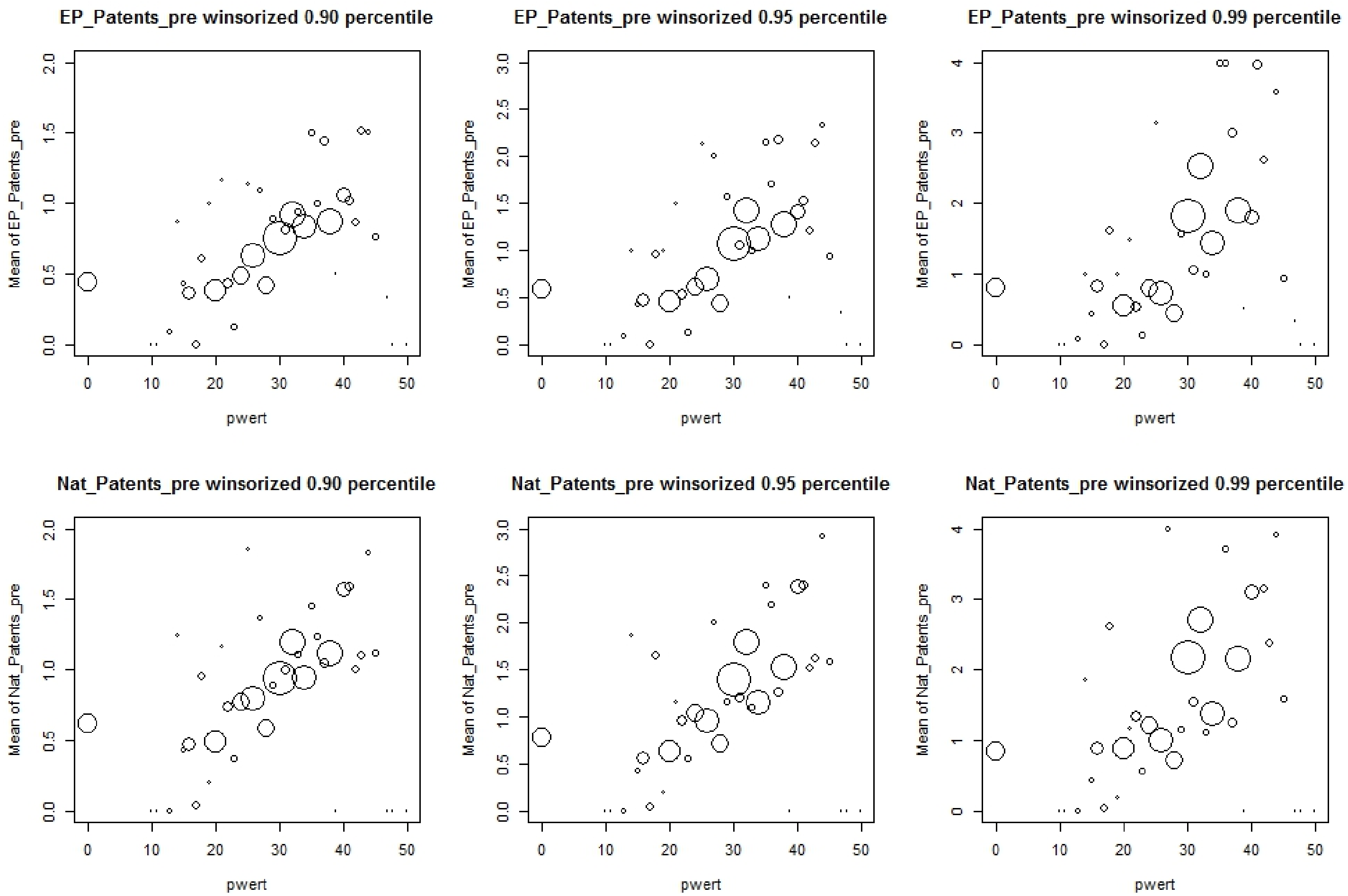
Notes: Areas of circles are proportional to number of observations. See section 1.3 for definition of variables. I report the median instead of the mean for firm age and project cost because variable values are dispersed.

Figure 21: Average propensity to file a European Patent in the 4 years preceding the funding application and in the 4 years after funding application by *pwert* score (upper row) and by *twert* score (lower row)



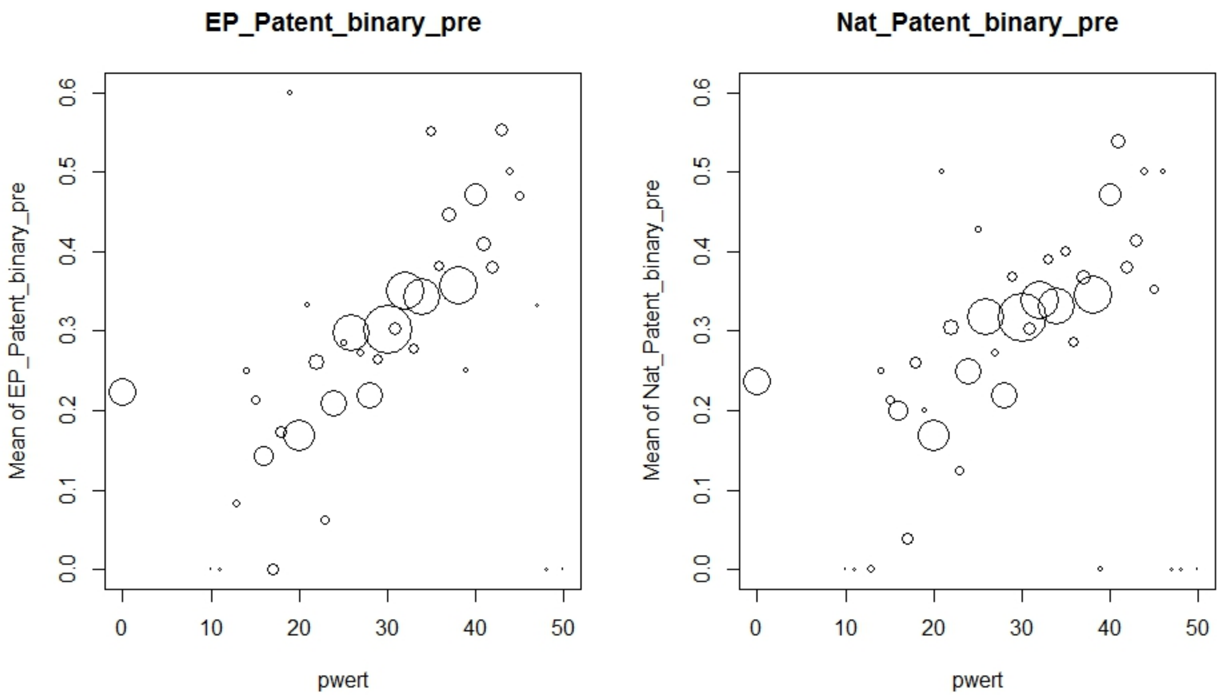
Note: Areas of circles are proportional to number of observations. Frequency of funding approval is plotted against *pwert* score in Figure 1 and against *twert* score in Figure 21.

Figure 23: Firm patent counts for 12 years preceding the funding application plotted against *pwert* score for European Patents (upper row) and National Patents (lower row)



Notes: Areas of circles are proportional to number of observations. See section 1.3 for definition of variables. Variables winsorized at 0.9 (left column), 0.95(middle column) and 0.99(right column) percentile.

Figure 24: Firm patenting propensities for 12 years preceding the funding application plotted against *pwert* score for European Patents (left) and National Patents (right)



Note: Areas of circles are proportional to number of observations. See section 1.3 for definition of variables.

Table 34: Validity check: applying the IV cubic spline model to project cost, firm age and firm sector controls

	(1) project cost	(2) project cost winsorized p=0.9	(3) firm age	(4) firm age winsorized p=0.9	(5) firm age> $5y$	(6) share manufacturing firms
approved	-2.895	-0.470	5.055**	2.179	0.0697	0.0849
	(1.877)	(0.421)	(2.360)	(1.361)	(0.0676)	(0.0666)
Observations	2,619	2,619	2,619	2,619	2,619	2,619
R-squared	0.035	0.078	0.011	0.011	0.008	0.019

Note: Clustered standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. In columns (2) and (4) dependent variable is winsorized at 0.9 percentile. In column (6), the dependent variable is the share of firms in sector Manufacturing of Metal Products, Electronics and Chemicals.

Table 35: Validity check: applying the IV cubic spline model to firm patenting history controls

	(1) EP Patents pre	(2) EP Patents pre winsorized p=0.9	(3) EP Patent binary pre	(4) Nat Patents pre	(5) Nat Patents pre winsorized p=0.9	(6) Nat Patent binary pre
approved	0.642	0.180	0.0537	-1.663	0.245	0.0857
	(1.249)	(0.174)	(0.0589)	(4.069)	(0.220)	(0.0599)
Observations	2,619	2,619	2,619	2,619	2,619	2,619
R-squared	0.013	0.031	0.033		0.023	0.021

Note: Clustered standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. In columns (2) and (5) dependent variable is winsorized at 0.9 percentile. Definition of variables can be found in section 1.3.3.

Table 36: Placebo regression for result on heterogeneous effect in treatment by firm age (Table 6)

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: EP Patent binary lagged by 4 years						
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
	IV: Cubic Splines	IV: 2 jumps	IV: 3 jumps	IV: Cubic Splines	IV: 2 jumps	IV: 3 jumps
	age$\leq 5y$	age$\leq 5y$	age$\leq 5y$	age$> 5y$	age$> 5y$	age$> 5y$
approved	-0.0311	-0.0247	0.0114	0.0121	0.00272	0.00509
	(0.0673)	(0.0622)	(0.0598)	(0.0616)	(0.0560)	(0.0557)
Observations	1,361	1,361	1,361	1,258	1,258	1,258
R-squared	0.273	0.274	0.277	0.382	0.382	0.382

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Full set of controls included in all models, but not reported. All patent outcomes and controls lagged by 4 years. Firms that are missing the year of incorporation counted as firms of age 0.

Robustness checks for results from section 1.5.1

In this section, I present robustness checks for the results discussed in section 1.5.1, the main result on the effect of funding approval on the propensity to file a European Patent and the result on heterogeneity by firm age, which are presented in Table 3 and Table 6.

I first consider two alternative specifications for controlling for the dependence of project quality on *pwert* score, illustrated in the middle column and the right column of Figure 2. As discussed in section 1.3, applications that received a *pwert* equal 0 are slightly ambiguous. Receiving a score of 0 means that the application failed a "knock-out" criterion. Table 37 and Table 45 show the results when I exclude such applications. In the second robustness check, in addition to excluding applications that received *pwert* equal 0, I control for the dependence of project quality on *pwert* score linearly, instead of using linear splines (Table 38 and Table 46). I also show the corresponding robustness checks for the specification when I extend the time frame over which patent outcomes are measured by an additional year (Table 39 and Table 40). As explained in section 1.5.1, firms for which the year of incorporation is missing in all likelihood never incorporated and are assigned an age of 0 in the specification presented in Table 6. In Table 47, I show that discarding such firms and splitting the sample at the new median age of 6 years produces very similar results.

I perform a number of additional robustness checks for the main result on the effect of funding approval on the propensity to file a European Patent: first, I restrict the sample to application that either fall in the point-range of the *pwert* score immediately before the steep increase in the funding approval probability or in the point-range after the steep increase (Table 41). Second, to account for the uncertainty that the break points of linear splines and the locations of jumps are chosen by best fit from data, I implement a bootstrap in which I re-sample firms with replacement and then choose break points and jumps by best fit for the obtained sample. Table 42 shows that the 95-percent confidence interval does not include 0 any model. In Table 43 and Table 44, I present the results of robustness checks that utilize the other evaluations scores of the agency (described in section 1.3). As discussed in section 1.3 footnote 7 and section 1.4 footnote 19, the primary score used in the analysis, the *pwert*, plays a distinctively important role in the grant approval process and produces the most pronounced discontinuities in the dependence of funding approval on evaluation score. In addition, the *pwert* score is independent of the identity of the firm and is calculated for each application separately. However, Table 43 shows that when I use the highly correlated *twert* score instead, which is an evaluation of the technical capabilities of the applicant firm, or in conjunction with the *pwert*, I find similar results. The remaining scores for the commercial evaluation of the applicant firm and proposed project, *wwert* and *fwert*, are calculated differently depending on the type of applicant firm and not necessarily comparable across applicants. Table 44 shows the results when I use the *pwert* score to construct my instrument and include all other scores as controls. Due to multicollinearity, the obtained estimate is less precise than the estimate in the main specification.

Table 37: Robustness check: excluding applications with *pwert* equal 0

Dependent variable: EP Patent binary post						
	(1) Model 1 IV: Cubic Splines	(2) Model 1 IV: Cubic Splines	(3) Model 2 IV: 2 jumps	(4) Model 2 IV: 2 jumps	(5) Model 3 IV: 3 jumps	(6) Model 3 IV: 3 jumps
approved	0.128**	0.122**	0.0876	0.107*	0.154**	0.124**
	(0.0634)	(0.0526)	(0.0718)	(0.0598)	(0.0693)	(0.0577)
Controls	NO	YES	NO	YES	NO	YES
Observations	2,480	2,480	2,480	2,480	2,480	2,480
R-squared	0.046	0.334	0.047	0.336	0.045	0.334

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. See section 1.3 for a discussion of why application with *pwert* score equal 0 are ambiguous.

Table 38: Robustness check: controlling for *pwert* score linearly (instead of using linear splines), applications with *pwert* equal 0 excluded

	Dependent variable: EP_Patent_binary_post					
	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 1	Model 2	Model 2	Model 3	Model 3
	IV: Cubic Splines	IV: Cubic Splines	IV: 2 jumps	IV: 2 jumps	IV: 3 jumps	IV: 3 jumps
	no controls	with controls	no controls	with controls	no controls	with controls
approved	0.0946	0.104**	0.0435	0.0816	0.0933	0.0952*
	(0.0625)	(0.0521)	(0.0631)	(0.0522)	(0.0620)	(0.0514)
pwert	0.00922***	0.000559	0.0113***	0.00146	0.00928***	0.000922
	(0.00317)	(0.00264)	(0.00317)	(0.00261)	(0.00315)	(0.00260)
Controls	NO	YES	NO	YES	NO	YES
Observations	2,480	2,480	2,480	2,480	2,480	2,480
R-squared	0.046	0.336	0.044	0.337	0.046	0.336

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 39: Robustness check: excluding applications with pwert equal 0, outcomes measured for one additional year

	Dependent variable: EP_Patent_binary_post measured for year of the funding application plus subsequent 4 years					
	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 1	Model 2	Model 2	Model 3	Model 3
	IV: Cubic Splines	IV: Cubic Splines	IV: 2 jumps	IV: 2 jumps	IV: 3 jumps	IV: 3 jumps
	no controls	with controls	no controls	with controls	no controls	with controls
approved	0.148**	0.139***	0.0994	0.118*	0.188***	0.154***
	(0.0652)	(0.0537)	(0.0739)	(0.0606)	(0.0715)	(0.0590)
Controls	NO	YES	NO	YES	NO	YES
Observations	2,480	2,480	2,480	2,480	2,480	2,480
R-squared	0.051	0.345	0.052	0.347	0.048	0.343

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 40: Robustness check: controlling for score linearly (instead using of linear splines), outcomes measured for one additional year, applications with pwert equal 0 excluded

Dependent variable: EP Patent binary post measured for year of the funding application plus subsequent 4 years						
	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 1	Model 2	Model 2	Model 3	Model 3
	IV: Cubic Splines	IV: Cubic Splines	IV: 2 jumps	IV: 2 jumps	IV: 3 jumps	IV: 3 jumps
approved	0.111*	0.116**	0.0512	0.0873*	0.116*	0.114**
	(0.0643)	(0.0533)	(0.0648)	(0.0531)	(0.0639)	(0.0526)
Controls	NO	YES	NO	YES	NO	YES
Observations	2,480	2,480	2,480	2,480	2,480	2,480
R-squared	0.051	0.346	0.049	0.348	0.051	0.347

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 41: Robustness check: using only application before (pwert $[19, 22]$) and after (pwert $[29, 30]$) the steep increase in the funding approval probability

Dependent variable: EP Patent binary post		
	(1) EP Patent binary post IV: pwert$\in [29, 30]$	(2) EP Patent binary lagged 4y IV: pwert$\in [29, 30]$ Placebo Test
approved	0.104***	0.00929
	(0.0336)	(0.0322)
project cost (in 100K Euro)	0.00362	0.00746
	(0.00502)	(0.00468)
project cost sqr	2.58e-05	-1.43e-06
	(8.82e-05)	(9.66e-05)
EP Patent binary pre	0.197***	0.231***
	(0.0586)	(0.0774)
EP Patents pre	0.0232**	0.0265**
	(0.00924)	(0.0133)
EP Patents pre sqr	-0.000315***	-0.000362**
	(9.91e-05)	(0.000153)
Nat Patent binary pre	0.103*	0.238***
	(0.0553)	(0.0671)
Nat Patents pre	0.00592	-0.00291
	(0.00731)	(0.00863)
Nat Patents pre sqr	5.47e-05	6.26e-05
	(6.00e-05)	(5.98e-05)
Constant	0.0371	0.144***
	(0.0473)	(0.0486)
Year FE, Sector FE, Age FE	YES	YES
Observations	689	689
R-squared	0.360	0.387

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. I report the intercept for a firm in sector Manufacturing of Metal Products, Electronics and Chemicals, Age$\geq$ 15 years, applying in 2005. Patent controls in column 2 are lagged by 4 years. Note that no controls for *pwert* score are included.

Table 42: Confidence intervals for the effect of funding approval on the propensity to file a European Patent from Bootstrap N=1000

	Dependent variable: EP Patent binary post					
	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 1	Model 2	Model 2	Model 3	Model 3
	IV: Cubic Splines	IV: Cubic Splines	IV: 2 jumps	IV: 2 jumps	IV: 3 jumps	IV: 3 jumps
approved						
90% confidence interval	[0.037;0.224]	[0.033;0.186]	[0.052;0.229]	[0.032;0.176]	[0.063;0.234]	[0.037;0.179]
95% confidence interval	[0.025;0.243]	[0.019;0.201]	[0.037;0.244]	[0.021;0.188]	[0.050;0.255]	[0.024;0.194]
99% confidence interval	[-0.002;0.270]	[-0.010;0.222]	[0.004;0.279]	[-0.006;0.215]	[0.027;0.281]	[0.004;0.217]
Controls	NO	YES	NO	YES	NO	YES

Note: Bootstrap Procedure: first clusters(=firms) are sampled, location of splines and jumps chosen by best fit from the first stage regression, then IV model from section 1.4 is re-estimated

Table 43: Robustness check: using the evaluation score *twert* as instrument

Dependent variable: EP Patent binary post						
	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 1	Model 2	Model 2	Model 3	Model 3
	IV: Cubic Splines	IV: Cubic Splines	IV: 2 jumps	IV: 2 jumps	IV: 3 jumps	IV: 3 jumps
	twert	twert,pwert jointly	twert	twert,pwert jointly	twert	twert,pwert jointly
approved	0.104*	0.107**	0.0995**	0.101**	0.110**	0.112***
	(0.0543)	(0.0435)	(0.0487)	(0.0401)	(0.0482)	(0.0396)
twert	-0.00120	0.000720	-0.000944	0.000672	-0.00124	0.000471
	(0.00179)	(0.00175)	(0.00182)	(0.00161)	(0.00180)	(0.00161)
linear spline twert 21	0.00168	-0.000729				
	(0.00274)	(0.00281)				
linear spline twert 24			0.00179	-0.000555	0.00176	-0.000524
			(0.00249)	(0.00276)	(0.00249)	(0.00276)
pwert		-0.00339*		-0.00298*		-0.00313*
		(0.00175)		(0.00168)		(0.00168)
linear spline pwert 21		0.00445				
		(0.00294)				
linear spline pwert 23				0.00442		0.00427
				(0.00289)		(0.00290)
Observations	2,617	2,617	2,617	2,617	2,617	2,617
R-squared	0.337	0.337	0.337	0.338	0.336	0.337

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Models (1) and (2) were estimated with cubic spline starting at $twert = 21$, models (3) and (4) were estimated with jumps at $twert = 24, 28$, models (5) and (6) were estimated with jumps at $twert = 24, 26, 28$. Full set of controls included, but not reported.

Table 44: Robustness check: including *twert*, *wwert* and *fwert* scores as controls, subset of applications where all scores are available.

	Dependent variable: EP Patent binary post					
	(1) Model 1 IV: Cubic Splines	(2) Model 1 IV: Cubic Splines	(3) Model 2 IV: 2 jumps	(4) Model 2 IV: 2 jumps	(5) Model 3 IV: 3 jumps	(6) Model 3 IV: 3 jumps
	all scores avail.	all scores avail.	all scores avail.	all scores avail.	all scores avail.	all scores avail.
approved	0.106**	0.0951*	0.0993**	0.0923**	0.108**	0.100**
	(0.0466)	(0.0514)	(0.0431)	(0.0467)	(0.0425)	(0.0463)
pwert	-0.00267	-0.00369**	-0.00228	-0.00338*	-0.00251	-0.00345*
	(0.00182)	(0.00188)	(0.00182)	(0.00177)	(0.00180)	(0.00177)
linear spline pwert 21	0.00366	0.00442				
	(0.00296)	(0.00313)				
linear spline pwert 23			0.00375	0.00444	0.00366	0.00433
			(0.00274)	(0.00285)	(0.00275)	(0.00286)
fwert		0.000919		0.000916		0.000888
		(0.00121)		(0.00121)		(0.00121)
wwert		0.00130		0.00139		0.00128
		(0.00172)		(0.00169)		(0.00169)
twert		7.81e-05		0.000115		2.11e-05
		(0.00132)		(0.00130)		(0.00130)
Observations	2,563	2,563	2,563	2,563	2,563	2,563
R-squared	0.339	0.340	0.339	0.341	0.339	0.340

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Full set of controls included , but not reported. Subset of applications where all evaluation scores are available. Columns (1), (3) and (5) report the estimates of the main specification (without controls for other scores) on this subset.

Table 45: Robustness check for heterogeneous effect in treatment by firm age: exluding applications with *pwert* equal 0

Dependent variable: EP_Patent_binary_post						
	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
	IV: Cubic Splines	IV: 2 jumps	IV: 3 jumps	IV: Cubic Splines	IV: 2 jumps	IV: 3 jumps
	age$\leq 5y$	age$\leq 5y$	age$\leq 5y$	age$> 5y$	age$> 5y$	age$> 5y$
approved	0.0457	0.0388	0.0486	0.210***	0.191**	0.209**
	(0.0743)	(0.0836)	(0.0787)	(0.0812)	(0.0843)	(0.0824)
Observations	1,274	1,274	1,274	1,206	1,206	1,206
R-squared	0.333	0.333	0.333	0.337	0.340	0.337

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Full set of controls included, but not reported. Firms that are missing the year of incorporation counted as firms of age 0. See discussion in section 1.3 for why applications that received *pwert* equal 0 are ambiguous.

Table 46: Robustness check for heterogeneous effect in treatment by firm age: controlling for score linearly (instead using of linear splines), applications with *pwert* equal 0 excluded

Dependent variable: EP_Patent_binary_post						
	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
	IV: Cubic Splines	IV: 2 jumps	IV: 3 jumps	IV: Cubic Splines	IV: 2 jumps	IV: 3 jumps
	age$\leq 5y$	age$\leq 5y$	age$\leq 5y$	age$> 5y$	age$> 5y$	age$> 5y$
approved	0.0413	0.0317	0.0397	0.169**	0.134*	0.148**
	(0.0705)	(0.0764)	(0.0747)	(0.0743)	(0.0704)	(0.0699)
Observations	1,274	1,274	1,274	1,206	1,206	1,206
R-squared	0.333	0.333	0.333	0.342	0.345	0.344

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Full set of controls included, but not reported. Firms that are missing the year of incorporation counted as firms of age 0.

Table 47: Heterogeneity in treatment by firm age, subset of applications by firms with available year of incorporation (2444 out of 2619 applications)

Dependent variable: EP Patent binary post						
	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
	IV: Cubic Splines	IV: 2 jumps	IV: 3 jumps	IV: Cubic Splines	IV: 2 jumps	IV: 3 jumps
	age$\leq 6y$	age$\leq 6y$	age$\leq 6y$	age$> 6y$	age$> 6y$	age$> 6y$
approved	0.0493	0.0616	0.0566	0.188***	0.166***	0.177***
	(0.0688)	(0.0643)	(0.0626)	(0.0689)	(0.0629)	(0.0620)
Observations	1,288	1,288	1,288	1,156	1,156	1,156
R-squared	0.322	0.322	0.322	0.344	0.347	0.345

Note: Cluster robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Full set of controls included, but not reported. In this subset of applications, the median firm age is 6 years.

Comparison of undirected R&D tax credit and directed research grants

In this section, I consider the model from section 1.6 and compare the research policy described in Proposition 3 with an alternative innovation policy instrument, an undirected R&D tax credit that uniformly lowers the cost of research.

Proposition 5. *Suppose that Assumptions 1 and 2 hold. Consider an R&D tax credit with a matching rate τ_{TC} that uniformly lowers the cost of all research projects for all firms by $(1 - \tau_{TC})$ percent. The remaining τ_{TC} percent of the cost are borne by the public.*

The following is true:

- *the threshold age of the current research line for exploration in the firm decision problem $\bar{T}$ weakly decreases in the matching rate of the R&D tax credit τ_{TC}*
- *the social planner can implement the socially optimal research policy through a R&D tax credit*
- *the payoff of the firm strictly increases in the matching rate of the R&D tax credit τ_{TC}*
- *the R&D tax credit that implements the socially optimal research policy at the lowest cost is strictly more costly to the public than the research grant policy from Proposition 3, unless socially and privately optimal research policies coincide*

Proof. See Appendix. □

Lowering the cost of research uniformly with an R&D tax credit encourages explorative research.[82] The R&D tax credit policy that implements the socially optimal research policy is more costly to the public than the research grant policy described in Proposition 3. In the next section, I solve the model explicitly for the case of one incremental follow-up project and

[82] To gain a better understanding of the scope and generality of this result, it is useful to relate it to existing models of exploration and search. First, consider a standard search problem where a decision maker samples i.i.d. draws from an arbitrary distribution at cost c. In such a model, a reduction in search cost c induces the DM to search longer. Now, suppose that instead of sampling i.i.d draws, the DM has alternative "projects" that she can research that differ in their risk/payoff profiles. Again, the decision maker searches for the best alternative. To fix ideas, suppose that each of the n projects $(P_j)_{j \in \{1,..,n\}}$ costs c_j to pursue, gives a payoff of X_j with probability p_j and 0 otherwise. Without loss of generality, suppose that the projects do not dominate each other so that $X_j > X_k$ implies $p_j < p_k$. The optimal policy in this problem was famously described by Weitzmann (1979): the DM calculates the reservation values of all n projects, and starts searching the projects in the order of their reservation values. When the DM finds a payoff that exceeds the reservation

show that the cost disadvantage of the R&D tax credit may be substantial. The reason is that the tax credit suffers from a double inefficiency in encouraging explorative research: not only are incremental projects subsidized, but also explorative research projects are overfunded.

The Model with one Incremental Follow-up Project per Research Line

In this section, I solve the model with the smallest possible state space. The set of Bellman equations that govern the solution are presented below, where I have already made use of the fact that exploration must be optimal for the young firm.

$$V_{Young} = V_{Estbl}^{obs} = -c_e + p_e(K + \beta V_{Estbl}^1) + (1 - p_e)\beta V_{Young} \tag{56}$$

$$V_{Estbl}^1 = max\Big\{K\lambda - c_i + \beta V_{Estbl}^{obs}; -c_e + p_e(K + \beta V_{Estbl}^1) + (1 - p_e)(K\lambda - c_i + \beta V_{Estbl}^{obs})\Big\} \tag{57}$$

---

value of all remaining unsearched projects she stops. The reservation values $(R_j)_{i\in\{1,..,n\}}$ solve

$$p_j(X_j - R_j) = c_j$$

and are given by

$$R_j = X_j - \frac{c_j}{p_j}$$

The reservation value has the following meaning: unless the decision maker already possesses a safe reward of value R_j, it is beneficial to explore project P_j. It is straightforward to see from the definition of the reservation value that reducing the search cost by a constant for all projects alters the *direction* of search. Since

$$\frac{d(R_j)}{dc_j} = -\frac{1}{p_j}$$

it is apparent that the reservation values of riskier projects are disproportionally affected by the reduction in cost. Since the cost of failure is lowered, the DM is more inclined to try riskier projects. More, generally if the rewards of the projects are distributed according to the CDFs $(F_j)_{i\in\{1,..,n\}}$, then

$$\frac{d(R_j)}{dc_j} = -\frac{1}{1 - F_j(R_j)}$$

Hence, riskier projects are researched systematically earlier. A tax credit with a matching rate τ affects the reservation values as follows:

$$\frac{d(R_j)}{d\tau} = -\frac{c_j}{p_j}$$

If costs are reduced proportionally by $(1 - \tau)$ percent for all projects, riskier projects are researched systematically earlier as long as costs are close to homogeneous (or if riskier project are more expensive). In the model presented in this paper, this requirement is embodied in Assumption 2. If costs are fully subsidized and therefore of no concern at all to the DM, she will attempt the project that promises the maximal reward irrespective of risk, simply because she does not care about the costs sunk in a failed attempt. There is no analogous result for the standard multi-bandit model. In the standard multi-bandit model, a proportional reduction in cost translates into uniform increases in the indices of all bandits (if costs are homogeneous). Hence, the effect of a reduction in the cost of research depends on how the innovation process is modelled.

It is not clear a priori whether exploitation is optimal for the firm that has the incremental follow-up project available. In this section, I refer to the firm that has the incremental follow-up project available as the "established firm", and to the firm that does not as the "young firm". The optimal policy is summarized below.

Proposition 6. *Consider the model with one incremental follow-up project and suppose that Assumptions 1 and 2 hold. Then, exploration is privately optimal for the established firm that has the incremental follow-up project available if and only if*

$$\frac{c_e}{p_e} - c_i - \frac{\beta(c_e - c_i)}{1 + p_e\beta} \leq K - (1-\beta)K\lambda - \beta\frac{p_e(K + \beta K\lambda)}{1 + p_e\beta} \tag{58}$$

Exploration is optimal for a social planner that has the incremental follow up project available if and only if

$$\frac{c_e}{p_e} - c_i - \frac{\beta(c_e - c_i)}{1 + p_e\beta} \leq SK - (1-\beta)SK\lambda - \beta\frac{p_e(SK + \beta SK\lambda)}{1 + p_e\beta} \tag{59}$$

In particular, if exploration is privately optimal for the established firm, then it must also be socially optimal. If exploration is socially, but not privately optimal, then a research grant with a matching rate τ_{RG}, defined as the solution to

$$\frac{c_e}{p_e}\Big(1 - \tau_{RG}\Big) - c_i = K - (1-\beta)K\lambda - \beta\frac{p_e(K + \beta K\lambda)}{1 + p_e\beta} + \frac{\beta(c_e - c_i)}{1 + p_e\beta} \tag{60}$$

implements the socially optimal research policy.

Proof. See Appendix. □

The research grant policy described in Proposition 5 not only discriminates between explorative and incremental projects, but also between established firms and young firms. I compare the research grant policy described in Proposition 5 with two alternative innovation policy instruments: a R&D tax credit that lowers the cost for all projects and firms uniformly, and a suboptimal simple research grant policy that only discriminates between explorative and incremental projects. Suppose from here on, that exploration is socially optimal, but not privately optimal for the established firm that has the incremental follow-up project available.

Proposition 7. *Consider the model with one incremental follow-up project, suppose that Assumptions 1 and 2 hold and assume that $c_i > 0$. Furthermore, suppose that exploration is* *socially optimal, but not privately optimal for the established firm that has the incremental follow-up project available.*

Let τ_{TC} *denote the matching rate of the most cost-effective R&D tax credit that implements the socially optimal research policy, let* τ'_{RG} *denote the matching rate of the most cost-effective simple research grant policy (defined above) that implements the socially optimal research policy, and consider the matching rate of the optimal research grant policy* τ_{RG} *defined in Proposition 5.*

Then, it must be true that

$$\tau_{TC} > \tau'_{RG} > \tau_{RG}$$

Proof. See Appendix. □

Proposition 6 clarifies that the R&D tax credit policy is expensive to the public not only because incremental research is subsidized, but also because explorative research projects need to be overfunded. There is a robust and simple intuition: suppose by contradiction that $\tau_{TC} \leq \tau'_{RG}$. By definition, the research grant with matching rate τ'_{RG} for the explorative project keeps the firm indifferent between the explorative and the incremental project. Thus, since the tax credit also funds the incremental project, it must be that the firm strictly prefers the incremental project if it is offered a tax credit with matching rate τ_{TC} instead. This double inefficiency results in a massive cost advantage for research grant policies. For example, at the parameter values $p_e = 0.6$, $c_i = c_e = 1$ and $\beta = 0.9$, the R&D tax credit is 2.01 times as expensive to the public as the simple research grant policy and 5 times as expensive as the optimal research grant policy.[83] It is worth noting that the payoff of the firm is highest under the tax credit policy.[84]

[83] I set $K = 1.7$, $\lambda = 0.7$ and $S = 2$ to make exploration socially, but not privately optimal for the established firm. The expected per-period cost to the public is $\tau_{TC} c_e + p_e(1-p_e)\tau_{TC} c_i$ for the R&D tax credit, $\tau'_{RG} c_e$ for the simple research grant policy and $p_e \tau_{RG} c_e$ for the optimal research grant policy. The ratio of the expected per-period cost of the R&D tax credit and the simple research grant policy is

$$\frac{\tau_{TC} c_e + p_e(1-p_e)\tau_{TC} c_i}{\tau'_{RG} c_e} = \frac{\tau_{TC}}{\tau'_{RG}}\left(1 + p_e(1-p_e)\frac{c_i}{c_e}\right) = \frac{1 + p_e(1-p_e)\frac{c_i}{c_e}}{1 - \frac{c_i}{c_e} p_e(1 - \beta(1-p_e))}$$

See the proof of proposition 6 for an expression of $\frac{\tau_{TC}}{\tau'_{RG}}$ in terms of parameters.

[84] Since all innovation policy instruments implement the same research policy and differ only in the payments to the firm.

Proofs of Propositions 1, 2, 4, 5 and 6

Proof of Proposition 1

For the first part, consider any pair V_{Estbl}^{T} and $V_{Estbl}^{\tilde{T}}$ and suppose that $T < \tilde{T}$. Consider the optimal policy starting at $\tilde{T}$. If the firm followed the same policy starting at T instead, then the incurred payoffs would be strictly higher until the research line would be renewed for the first time (since the stage payoffs of incremental projects are strictly higher for research lines of younger age) and exactly equal ever after. Since there is a non-zero chance that the research line is not renewed immediately, this shows that the value of the optimal policy at T must be strictly higher than the value of the optimal policy at $\tilde{T}$. The same argument holds true when comparing V_{Young} or V_{Estbl}^{Obs} with any V_{Estbl}^{T}. To see that $V_{Young} = V_{Estbl}^{Obs}$ note that they obey the same Bellman equation. For the second part, consider inequality (5). Since V_{Estbl}^{T} and $K\lambda^{T}$ are decreasing in T, it must be that the gain from renewing the research line increases as T increases. For the third part, suppose by contradiction that the optimal strategy is to not undertake the project. Then, the value of a young firm must be 0. However, the payoff of the (non-stationary) strategy of trying to explore exactly once and then exploit the research line until it becomes obsolete is strictly positive by Assumption 1. This shows that undertaking the research project must be optimal for a young firm.

Proof of Proposition 2

The argument is split up into multiple steps. Throughout the proof, I refer to the age of the research line as the "state", with states $1, 2, .., N$, and refer to a firm without an active research line, i.e. young firms and firms with an obsolete research line, as being in state $N+1$. To reiterate, the social planner problem is the firm decision problem with all research payoffs K,$K\lambda$,$K\lambda^2$, .. scaled up by the constant factor $S > 1$.

I know from Proposition 1 that there are only $N + 1$ candidates for the optimal policy in the firm decision problem and only $N + 1$ candidates for the optimal policy in the social planner problem. The $N + 1$ candidates for the optimal policy are

1. Exploration is optimal in states $T \geq 1$
2. Exploration is optimal in states $T \geq 2$, exploitation is optimal in state $T = 1$
3. Exploration is optimal in states $T \geq 3$, exploitation is optimal in states $T < 3$

and so on. I denote these policies by $\pi^1, \pi^2, \pi^3, .., \pi^{N+1}$ with the indices given by the first state at which exploration happens. Furthermore, I let V_n^S denote the value of policy π^S in state n and denote the value function function of policy π^S by V^S.

Since the state-space is finite, one way to single out the optimal policy is to do the following: first I compute V^{N+1}, the value function of policy π^{N+1}. If the Bellman equations are satisfied in all states $T \geq 1$ for value function V^{N+1}, I stop and conclude that policy π^{N+1} is indeed optimal. If there is any state in which the Bellman equation is not satisfied, I eliminate π^{N+1} as a candidate policy and note that the optimal policy must be in the set $\{\pi^N, \pi^{N-1}, \pi^{N-2}, .., \pi^1\}$. I proceed by computing V^N, the value of policy π^N. Again, if all Bellman equations are satisified I stop, otherwise I eliminate π^N and note that the optimal policy must be exploring earlier and lie in $\{\pi^{N-1}, \pi^{N-2}, .., \pi^1\}$. I proceed with candidate policy π^{N-1} and so on.

The argument that proves the statement is the following: in this algorithm, if π^T can be eliminated as a candidate for the optimal policy in the firm decision problem (implying that the firm starts exploring earlier than T), it can also be eliminated as a candidate in the social planner problem. This implies that exploration happens (weakly) earlier in the social planner problem, and for a (weakly) larger number of states. I will state and prove two Lemmas that are jointly sufficient for the result. I repeat the argument after the proof of Lemma 2.

Lemma 1. *If for policy π^T, the Bellman equation is violated in any state $t < T$ in the firm decision problem, it must also be violated in the social planner problem.*

Proof. I start by computing the value of policy π^T in state T, V_T^T, assuming that $T < N+1$. Policy π^{N+1} will be dealt with at the end of this proof. I denote by s the integer that satisfies $T + s = N$. By definition, the value function V^T satisfies the following set of equations

$$V_t^T = K\lambda^t - c_i + \beta V_{t+1}^T \text{ for all } t \text{ s.t. } 1 \leq t < T \tag{61}$$

and

$$V_{T+n}^T = -c_e + p_e(K + \beta V_1^T) + (1-p_e)(K\lambda^{T+n} - c_i + \beta V_{T+n+1}^T) \text{ for all } n \text{ s.t. } 0 \leq n \leq s \tag{62}$$

At state $N+1$ it holds that

$$V_{N+1}^T = -c_e + p_e(K + \beta V_1^T) + (1-p_e)(\beta V_{N+1}^T) \tag{63}$$

I now proceed by iteratively plugging equations (62) and (63) into each other to express V_T^T solely as a function of the stage-payoffs, V_{N+1}^T and V_1^T. I obtain

$$\begin{aligned} V_T^T = (1-p_e)^{s+1}\beta^{s+1}V_{N+1}^T + &\sum_{t=0}^{s}(1-p_e)^t\beta^{t+1}p_eV_1^T \\ &- \sum_{t=0}^{s}(1-p_e)^{t+1}\beta^t c_i - \sum_{t=0}^{s}(1-p_e)^t\beta^t c_e + \Psi K \end{aligned} \tag{64}$$

To save on notation, I express the term that multiplies K simply as ΨK. If I make use of the facts that

$$V_{N+1}^T = \frac{p_e(K + \beta V_1^T) - c_e}{1-(1-p_e)\beta}$$

(obtained from equation (62)) and

$$V_1^T = \beta^{T-1}V_T^T + \sum_{t=1}^{T-1}\beta^{t-1}(K\tau^t - c_i)$$

(obtained from iteratively plugging the equations from (61) into each other), I can express V_T^T solely in terms of stage payoffs. (64) becomes

$$\begin{aligned} V_T^T = &\sum_{t=0}^{s}(1-p_e)^t\beta^t p_e(\beta^T V_T^T - \sum_{j=1}^{T-1}\beta^j c_i) + (1-p_e)^{s+1}\beta^{s+1}p_e\frac{\beta^T V_T^T - \sum_{j=1}^{T-1}\beta^j c_i}{1-(1-p_e)\beta} \\ &- \sum_{t=0}^{s}(1-p_e)^{t+1}\beta^t c_i - \sum_{t=0}^{s}(1-p_e)^t\beta^t c_e - \frac{(1-p_e)^{s+1}\beta^{s+1}c_e}{1-(1-p_e)\beta} + \Psi' K \end{aligned} \tag{65}$$

Bringing all V_T^T-terms to the LHS and expressing all geometric sums with the summation formula, this is

$$
\begin{aligned}
V_T^T \Big\{1 - p_e\beta^T\Big(\frac{1-\beta^{s+1}(1-p_e)^{s+1}}{1-(1-p_e)\beta} + \frac{\beta^{s+1}(1-p_e)^{s+1}}{1-(1-p_e)\beta}\Big)\Big\} = \\
- c_i\Big\{\Big(p_e\frac{\beta-\beta^T}{1-\beta}\Big)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta} + \Big(p_e\frac{\beta-\beta^T}{1-\beta}\Big)\frac{(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\Big\} \\
- c_i\Big\{(1-p_e)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\Big\} \\
- c_e\Big\{\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta} + \frac{(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\Big\} + \Psi' K
\end{aligned}
\tag{66}
$$

(66) simplifies to

$$
\begin{aligned}
V_T^T\Big(1 - \frac{p\beta^T}{1-(1-p_e)\beta}\Big) = \\
- c_i\Big\{\Big(p_e\frac{\beta-\beta^T}{1-\beta}\Big)\frac{1}{1-(1-p_e)\beta} + (1-p_e)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\Big\} \\
- c_e\frac{1}{1-(1-p_e)\beta} + \Psi' K
\end{aligned}
\tag{67}
$$

With this expression for V_T^T at hand, I can express the reservation value condition for state $t = T - n \leq T - 1$ as

$$
\begin{aligned}
\frac{c_e}{p_e} - c_i &\geq \beta(V_1^T - V_{T-(n-1)}^T) + K - K\lambda^{T-n} \\
&= \beta(\beta^{T-1}V_T^T - \sum_{j=1}^{T-1}\beta^{j-1}c_i - \beta^n V_T^T + \sum_{j=1}^{n}\beta^{j-1}c_i) + \Psi'' K \\
&= -c_i\frac{\beta^{n+1}-\beta^T}{1-\beta} + \frac{\beta(\beta^n - \beta^{T-1})}{1-\frac{p\beta^T}{1-(1-p_e)\beta}}\Big\{\frac{c_e}{1-(1-p_e)\beta} \\
&+ c_i\Big(\Big(p_e\frac{\beta-\beta^T}{1-\beta}\Big)\frac{1}{1-(1-p_e)\beta} + (1-p_e)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\Big)\Big\} + \Psi''' K
\end{aligned}
\tag{68}
$$

(68) simplifies to

$$
\frac{c_e}{p_e} - c_i - \frac{\beta(\beta^n - \beta^{T-1})}{1-\frac{p_e\beta^T}{1-(1-p_e)\beta}}\left\{\frac{c_e - c_i}{1-(1-p_e)\beta} + c_i(1-p_e)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta}\right\} \geq \Psi''' K \tag{69}
$$

I will now argue that the LHS expression of (69) is in fact strictly positive. Fix $c_e > 0$ and consider the LHS expression of (69) as a linear function of c_i on $[0, c_e/p_e]$.

First, suppose that $c_i = 0$. The LHS expression of (69) is then

$$\frac{c_e}{p_e} - \frac{\beta(\beta^n - \beta^{T-1})}{1 - \frac{p_e\beta^T}{1-(1-p_e)\beta}} \frac{c_e}{(1-(1-p_e)\beta} > \frac{c_e}{p_e} - \frac{\beta(1 - \beta^{T-1})}{1 - \frac{p_e\beta^T}{1-(1-p_e)\beta}} \frac{c_e}{(1-(1-p_e)\beta} =$$
$$= \frac{c_e}{p_e} - \frac{\beta(1-\beta^{T-1})c_e}{(1-(1-p_e)\beta - p_e\beta^T} = c_e \frac{1 - \beta + p_e\beta - p_e\beta^T - p_e\beta + p_e\beta^T}{p_e(1-(1-p_e)\beta - p_e\beta^T)} =$$
$$= c_e \frac{1-\beta}{p_e(1-(1-p_e)\beta - p_e\beta^T)} > 0$$

Next, suppose that $c_i = c_e/p_e$. The LHS expression of (69) becomes

$$-\frac{\beta(\beta^n - \beta^{T-1})}{1 - \frac{p_e\beta^T}{1-(1-p_e)\beta}} \left\{ \frac{c_e - \frac{c_e}{p_e}}{1-(1-p_e)\beta} + \frac{c_e}{p_e}(1-p_e)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta} \right\} =$$
$$= -\frac{\beta(\beta^n - \beta^{T-1})}{1 - \frac{p_e\beta^T}{1-(1-p_e)\beta}} \left\{ \frac{c_e - \frac{c_e}{p_e}}{1-(1-p_e)\beta} + \left(\frac{c_e}{p_e} - c_e\right)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta} \right\} =$$
$$= \frac{\beta(\beta^n - \beta^{T-1})}{1 - \frac{p_e\beta^T}{1-(1-p_e)\beta}} \left\{ \left(\frac{c_e}{p_e} - c_e\right)\frac{(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta} \right\} \geq 0$$

Since the LHS expression of (69) is linear in c_i, strictly positive at 0 and weakly positive at c_e/p_e, it must be strictly positive for all values in the interior of the interval $[0, c_e/p_e]$.

Now, suppose that reservation value condition (69) fails, meaning that the Bellman equation for the firm is not satisified for state $t = T - n$. But then, it must also be true that (69) fails for the social planner in state $T - n$, since, in this case

$$0 < \frac{c_e}{p_e} - c_i - \frac{\beta(\beta^n - \beta^{T-1})}{1 - \frac{p_e\beta^T}{1-(1-p_e)\beta}} \left\{ \frac{c_e - c_i}{1-(1-p_e)\beta} + c_i(1-p_e)\frac{1-(1-p_e)^{s+1}\beta^{s+1}}{1-(1-p_e)\beta} \right\} \tag{70}$$
$$< \Psi''' K < \Psi''' SK$$

This proves the Lemma for all policies π^T, $T < N + 1$. Consider therefore policy π^{N+1}. Repeating all steps from equations (61) to (69), it is straightforward to verify that the value at state $N+1$ for this policy satisfies

$$V_{N+1}^{N+1}\Big(1-\frac{p_e\beta^{N+1}}{1-(1-p_e)\beta}\Big)=-c_i\frac{p_e\sum_{t=1}^{N}\beta^t}{1-(1-p_e)\beta}-\frac{c_e}{1-(1-p_e)\beta}+\Psi K \tag{71}$$

The analogue to the reservation value condition (69) at state $N-n$ is

$$\frac{c_e}{p_e}-c_i-\frac{\beta(\beta^n-\beta^N)}{1-\frac{p_e\beta^{N+1}}{1-(1-p_e)\beta}}\Big(\frac{c_e-c_i}{1-(1-p_e)\beta}\Big)\geq\Psi' K \tag{72}$$

By the same argument as before, the LHS is strictly positive for all c_i and c_e s.t $c_i<c_e/p_e$. Thus, it must again be the case that if (72) fails in the firm decision problem, it also fails in the social planner problem, since in this case

$$0<\frac{c_e}{p_e}-c_i-\frac{\beta(\beta^n-\beta^N)}{1-\frac{p_e\beta^{N+1}}{1-(1-p_e)\beta}}\Big(\frac{c_e-c_i}{1-(1-p_e)\beta}\Big)<\Psi' K<\Psi' SK \tag{73}$$

This concludes the proof. □

Lemma 2. *Suppose that for policy π^T the Bellman equations of the firm decision problem are satisfied in all states $t<T$. Then, at least one policy in $\{\pi^T,\pi^{T+1},..,\pi^N,\pi^{N+1}\}$ must be optimal in the firm decision problem.*

Proof. Since the Bellman equation is satisfied in all states $t<T$, I know that

$$\begin{aligned}V_t^T=K\lambda^t-c_i+&\beta V_{t+1}^T\\&\geq c_e+p_e(K+\beta V_1^T)+(1-p_e)(K\lambda^t-c_i+\beta V_{t+1}^T)\text{ for all }t<T\end{aligned} \tag{74}$$

or equivalently

$$\frac{c_e}{p_e}-c_i\geq\beta(V_1^T-V_{t+1}^T)+K-K\lambda^t\text{ for all }t<T \tag{75}$$

I now iterate on the Bellman-operator, starting from value function V^T. Hence, the first iteration is given by

$$V_t^{T,1}=max\{K\lambda^t-c_i+\beta V_{t+1}^T;c_e+p_e(K+\beta V_1^T)+(1-p_e)(K\lambda^t-c_i+\beta V_{t+1}^T)\} \tag{76}$$

for all t. The subsequent iterations are given by

$$V_t^{T,n} = max\{K\lambda^t - c_i + \beta V_{t+1}^{T,n-1}; c_e + p_e(K + \beta V_1^{T,n-1}) + (1-p_e)(K\lambda^t - c_i + \beta V_{t+1}^{T,n-1})\} \tag{77}$$

for all t. Because $V^{T,1} \geq V^T$ by construction, it follows from the monotonicity of the Bellman operator that $(V^{T,n})_{n\in\mathbb{N}}$ is an increasing sequence. Since I am iterating on the Bellman-operator, the sequence converges to the value function of the optimal policy. I denote the weakly positive increase in value in iteration n at the fixed state T, $V_T^{T,n} - V_T^{T,n-1}$, by Δ_T^n.

Given $(\Delta_T^n)_{n\in\mathbb{N}}$, I first compute the implied value functions $V_t^{T,n}$ in states $t < T$ and for all subsequent iterations n. I compute the value functions for the first three iterations by hand and then prove the inductive step to establish the validity for all $n \in \mathbb{N}$. For the first iteration, I know that (76) implies that

$$V_t^{T,1} = V_t^T = K\lambda^t - c_i + \beta V_{t+1}^T \text{ for all } t < T \tag{78}$$

Also, by definition

$$V_T^{T,1} = V_T^T + \Delta_T^1 \tag{79}$$

Now, consider the second iteration. First, I have to determine whether exploitation is still optimal in all states $t < T$, given $V^{T,1}$. However, for all states $t < T-1$

$$\begin{aligned} &\beta(V_1^{T,1} - V_{t+1}^{T,1}) + K - K\lambda^t \\ &= \beta(V_1^T - V_{t+1}^T) + K - K\lambda^t \text{ (by (78))} \\ &\leq \frac{c_e}{p_e} - c_i \text{ (by (75))} \end{aligned}$$

Furthermore, for state $T-1$,

$$\begin{aligned} &\beta(V_1^{T,1} - V_T^{T,1}) + K - K\lambda^t \\ &= \beta(V_1^T - V_T^T - \Delta_T^1) + K - K\lambda^t \text{ (by (78) and (79))} \\ &\leq \beta(V_1^T - V_t^T) + K - K\lambda^t \\ &\leq \frac{c_e}{p_e} - c_i \text{ (by (75))} \end{aligned}$$

Thus, exploitation is still optimal in all states $t < T$. This implies that in states $t < T - 1$

$$\begin{aligned} V_t^{T,2} &= \beta V_{t+1}^{T,1} + K\lambda^t - c_i \\ &= \beta V_{t+1}^{T} + K\lambda^t - c_i \text{ (by(78))} \\ &= V_t^T \end{aligned}$$

and in state $T - 1$

$$\begin{aligned} V_{T-1}^{T,2} &= \beta V_T^{T,1} + K\lambda^t - c_i = \\ &= \beta(V_T^T + \Delta_T^1) + K\lambda^t - c_i = \\ &= \beta V_T^T + K\lambda^t - c_i + \beta\Delta_T^1 = \\ &= V_{T-1}^T + \beta\Delta_T^1 \end{aligned}$$

I record that $V_t^{T,2} = V_t^T$ for all $t < T - 1$ and $V_{T-1}^{T,2} = V_{T-1}^T + \beta\Delta_T^1$. By definition, $V_T^{T,2} = V_T^T + \Delta_T^1 + \Delta_T^2$. It is instructive to also compute the third iteration. Again, I have to determine whether exploration is still optimal in all states $t < T$, given $V_T^{T,2}$. For states $t < T - 2$,

$$\begin{aligned} &\beta(V_1^{T,2} - V_{t+1}^{T,2}) + K - K\lambda^t \\ &= \beta(V_1^T - V_{t+1}^T) + K - K\lambda^t \\ &\leq \frac{c_e}{p_e} - c_i \end{aligned}$$

For state $T - 2$,

$$\begin{aligned} &\beta(V_1^{T,2} - V_{T-1}^{T,2}) + K - K\lambda^{T-2} \\ &= \beta(V_1^T - V_{T-1}^T - \beta\Delta_T^1) + K - K\lambda^{T-2} \\ &\leq \beta(V_1^T - V_{T-1}^T) + K - K\lambda^{T-2} \\ &\leq \frac{c_e}{p_e} - c_i \end{aligned}$$

And for state $T-1$,

$$\begin{aligned} &\beta(V_1^{T,2} - V_T^{T,2}) + K - K\lambda^{T-1} \\ &= \beta(V_1^T - V_T^T - \Delta_T^1 - \Delta_T^2) + K - K\lambda^{T-1} \\ &\leq \beta(V_1^T - V_T^T) + K - K\lambda^{T-2} \\ &\leq \frac{c_e}{p_e} - c_i \end{aligned}$$

Thus, exploitation is still optimal in all states $t < T$. Thus, I have that

$$V_t^{T,3} = K\lambda^t - c_i + \beta V_{t+1}^{T,2} \text{ for all } t < T \tag{80}$$

Plugging in the expressions for $V_{t+1}^{T,2}$, I record that $V_t^{T,3} = V_t^T$ for $t < T-2$, $V_{T-2}^{T,3} = V_{T-2}^T + \beta^2\Delta_T^1$ and $V_{T-1}^{T,3} = V_{T-1}^T + \beta(\Delta_T^1 + \Delta_T^2)$. Again, by definition $V_T^{T,3} = V_T^T + \Delta_T^1 + \Delta_T^2 + \Delta_T^3$. If I keep expanding, I have that for state $t = T - s$, s.t. $0 \leq s \leq T-1$, in iteration n,

$$V_{T-s}^{T,n} = V_{T-s}^T + \beta^s \sum_{j=1}^{n-s} \Delta_T^j \tag{81}$$

I now prove the inductive step. Suppose (81) holds in iteration n. I will show that (81) holds for iteration $n+1$. First, I check whether in iteration $n+1$, exploitation is still optimal in all states $t < T$ given $V^{T,n}$. For states $t = T - s$, s.t. $1 \leq s \leq T-1$,

$$\begin{aligned} &\beta(V_1^{T,n} - V_{T-(s-1)}^{T,n}) + K - K\lambda^{T-s} = \\ &= \beta\Big(V_1^T + \beta^{T-1}\sum_{j=1}^{n-(T-1)} \Delta_T^j - V_{T-(s-1)}^T - \beta^{s-1}\sum_{j=1}^{n-(s-1)} \Delta_T^j\Big) + K - K\lambda^{T-s} = \\ &= \beta\Big(V_1^T - V_{T-(s-1)}^T\Big) + K - K\lambda^{T-s} + \beta\Big(\beta^{T-1}\sum_{j=1}^{n-(T-1)} \Delta_T^j - \beta^{s-1}\sum_{j=1}^{n-(s-1)} \Delta_T^j\Big) \\ &\leq \beta\big(V_1^T - V_{T-(s-1)}^T\big) + K - K\lambda^{T-s} \leq \frac{c_e}{p_e} - c_i \end{aligned}$$

Hence, exploitation is still optimal in states $t < T$. Thus, for states $t = T - s$, s.t. $1 \leq s \leq$

$T-1$,

$$V_{T-s}^{T,n+1} = \beta V_{T-(s-1)}^{T,n} + K\lambda^{T-s} - c_i =$$
$$= \beta\Big(V_{T-(s-1)}^{T} + \beta^{s-1}\sum_{j=1}^{n-(s-1)} \Delta_T^j\Big) + K\lambda^{T-s} - c_i =$$
$$= V_{T-s}^{T} + \beta^{s}\sum_{j=1}^{(n+1)-s} \Delta_T^j$$

and by definition

$$V_T^{T,n+1} = V_T^T + \sum_{j=1}^{n+1} \Delta_T^j$$

This establishes (81) for all steps of the iteration. Since $V^{T,n}$ converges to the value function of the optimal policy, call it V, it must be that for states $t = T-s$, s.t. $0 \le s \le T-1$,

$$V_{T-s} = lim_{n\to\infty} V_{T-s}^{T,n} = V_{T-s}^T + \beta^s \sum_{j=1}^{\infty} \Delta_T^j \tag{82}$$

Since $\sum_{j=1}^{\infty} \Delta_T^j$ is just $V_T - V_T^T$, this series is definitely finite. I am now in a position where I can determine whether exploitation is preferred at states $t < T$ under the optimal policy. First, suppose that $\sum_{j=1}^{\infty} \Delta_T^j > 0$. Note that in this case

$$\beta\Big(V_1 - V_T\Big) + K - K\lambda^T$$
$$= \beta\Big(V_1^T - V_T^T\Big) + \beta\Big(\beta^{T-1}\sum_{j=1}^{\infty} \Delta_T^j - \sum_{j=1}^{\infty} \Delta_T^j\Big) + K - K\lambda^T$$
$$< \beta\Big(V_1^T - V_T^T\Big) + K - K\lambda^T \le \frac{c_e}{p_e} - c_i$$

In this case, exploitation is strictly preferred at state $T-1$ under the optimal policy, which implies that the optimal policy must be an element of $\{\pi^T, \pi^{T+1}, .., \pi^N, \pi^{N+1}\}$. If $\sum_{j=1}^{\infty} \Delta_T^j = 0$, meaning that π^T is optimal, the claim is trivially true. □

Suppose that in the elimination algorithm described in the beginning of the proof, I have already eliminated $\{\pi^{T+1}, .., \pi^N, \pi^{N+1}\}$ as candidates for the optimal policies, for both the firm decision problem and the social planner problem. Hence, the optimal policies of the firm decision problem and the social planner problem must be elements of $\{\pi^1, .., \pi^T\}$. Suppose that all Bellman equations for π^T hold in the firm decision problem. Thus, π^T is the optimal policy in the firm decision problem and the algorithm stops. Since the optimal policy for the social planner is in $\{\pi^1, .., \pi^T\}$, it holds true that the social planner starts exploring weakly earlier than the firm. If at least one Bellman equation for π^T is violated in the firm decision problem, π^T is eliminated as a candidate for the optimal policy in the firm decision problem and the actual optimal policy lies in $\{\pi^1, .., \pi^{T-1}\}$. I will argue that it must be that one of the Bellman equations for the states $t < T$ is violated. By contradiction, suppose not. Then, by Lemma 2, this means that at least one element of $\{\pi^T, \pi^{T+1}, .., \pi^N, \pi^{N+1}\}$ is an optimal policy in the firm decision problem. However, those policies were already eliminated in the earlier rounds. Thus, it must be that at least one of the Bellman equations in some state $t < T$ is violated. Hence, by Lemma 1, this implies that the same Bellman equation is also violated in the social planner problem and π^T can be also eliminated as a candidate for the optimal policy in the social planner problem. This concludes the proof.

Proof of Proposition 4

Proposition 4 is a restatement of Proposition 2. As mentioned before, an equivalent way of looking at the social planner problem is to consider it as the firm decision problem with all research cost lowered uniformly by the factor $1/S$. To see the first claim, consider the firm decision problem at an R&D tax credit with matching rate τ^1_{TC}, and at an R&D tax credit with matching rate τ^2_{TC}, where $\tau^1_{TC} > \tau^2_{TC}$. Then, the decision problem at the higher matching rate corresponds to the decision problem at the lower matching with research costs lowered by factor $(1-\tau^1_{TC})/(1-\tau^2_{TC})$. Therefore, in analogy to Proposition 2, the threshold age for exploration must be lower in the decision problem at the higher matching rate. The second claim follows from the observation that, if the matching rate of the tax credit is chosen such that $(1-\tau_{TC}) = 1/S$, the firm has the same objective as the social planner. The third claim is obvious: since costs are uniformly lowered, the payoff of the firm must increase. The last claim follows from the following consideration: since both policies implement the same research policy, the firm incurs the same research rewards (i.e. the K, $K\lambda$ and so on) under both policies. However, the payoff of the firm is higher under the R&D tax credit policy than under the research grant policy. Therefore, it must be that a higher share of the overall costs are borne by the public.

Proof of Proposition 5

One straightforward way to check whether exploitation is optimal for the established firm is to compute the value function of the policy that exploits the incremental project, denote it by $\tilde{V}$, and check whether the established firm's Bellman equation is satisfied. Thus,

$$\tilde{V}^1_{Estbl} = K\lambda - c_i + \beta \tilde{V}^{obs}_{Estbl} = K\lambda - c_i + \beta \tilde{V}_{Young}$$

$$\tilde{V}_{Young} = \tilde{V}^{obs}_{Estbl} = -c_e + p_e(K + \beta \tilde{V}^1_{Estbl}) + (1 - p_e)\beta \tilde{V}_{Young}$$

imply

$$\tilde{V}_{Young} = \tilde{V}^{obs}_{Estbl} = \frac{p_e(K + \beta(K\lambda - c_i)) - c_e}{(1-\beta)(1 + p_e\beta)} \tag{83}$$

and

$$\tilde{V}^1_{Estbl} = \beta \frac{p_e(K + \beta(K\lambda - c_i)) - c_e}{(1-\beta)(1 + p_e\beta)} + K\lambda - c_i \tag{84}$$

Exploitation is indeed optimal for the established firm if and only if the reservation value condition

$$\frac{c_e}{p_e} - c_i \geq \beta(\tilde{V}^1_{Estbl} - \tilde{V}^{obs}_{Estbl}) + K - K\lambda \tag{85}$$

is satisfied. Plugging in the value function of the conjectured policy, this is equivalent to

$$\frac{c_e}{p_e} - c_i \geq \beta(K\lambda - c_i) + \beta(\beta - 1)\frac{p_e(K + \beta(K\lambda - c_i)) - c_e}{(1-\beta)(1 + p_e\beta)} + K - K\lambda \tag{86}$$

Inequality (86) simplifies to

$$\frac{c_e}{p_e} - c_i - \frac{\beta(c_e - c_i)}{1 + p_e\beta} \geq K - (1-\beta)K\lambda - \beta\frac{p_e(K + \beta K\lambda)}{1 + p_e\beta} \tag{87}$$

Exploitation is optimal for the established firm if and only if (87) is satisfied. If this is the case, the value function of the firm (for the optimal policy) is given by $V^1_{Estbl} = \tilde{V}^1_{Estbl}$ and $V_{Young} = V^{obs}_{Estbl} = \tilde{V}_{Young} = \tilde{V}^{obs}_{Estbl}$. If (87) is not satisfied, then exploration is optimal for the established firm. In this case, the value function can be obtained by solving the system of equations (9) and (10), using the fact that exploration is optimal for the established firm. Hence, exploration is optimal for the established firm if and only if

$$\frac{c_e}{p_e} - c_i - \frac{\beta(c_e - c_i)}{1 + p_e\beta} \leq K - (1-\beta)K\lambda - \beta\frac{p_e(K + \beta K\lambda)}{1 + p_e\beta} \tag{88}$$

holds. If inequality (88) holds for the established firm, it must also hold for the social planner since

$$
\begin{aligned}
0 < \frac{c_e}{p_e} - c_i - &\frac{\beta(c_e - c_i)}{1 + p_e\beta} \\
&\leq K - (1-\beta)K\lambda - \beta\frac{p_e(K + \beta K\lambda)}{1 + p_e\beta} \\
&< S\Big\{K - (1-\beta)K\lambda - \beta\frac{p_e(K + \beta K\lambda)}{1 + p_e\beta}\Big\} \\
&= SK - (1-\beta)SK\lambda - \beta\frac{p_e(SK + \beta SK\lambda)}{1 + p_e\beta}
\end{aligned}
\tag{89}
$$

A quick way to see that the expression

$$\frac{c_e}{p_e} - c_i - \frac{\beta(c_e - c_i)}{1 + p_e\beta}$$

is strictly positive is to fix $c_e > 0$ and consider the expression as a linear function of c_i on $[0, c_e/p_e]$. The function is strictly positive at $c_i = 0$, since

$$c_e(\frac{1}{p_e} - \frac{\beta}{1 + p_e\beta}) > 0$$

and strictly positive at $c_i = c_e/p_e$ because

$$\frac{\beta(\frac{c_e}{p_e} - c_e)}{1 + p_e\beta} > 0$$

Since the expression is linear in c_i and strictly positive on both endpoints of the interval $[0, c_e/p_e]$ it must be strictly positive for all values in the interval. The optimal research grant for the explorative project with matching rate τ_{RG} that induces exploration (when exploration is not privately optimal) makes (87) hold with equality. Note that the research grant only multiplies the c_e-term in the stage payoff, not in the continuation payoff.

Proposition 6

The lowest R&D tax credits τ_{TC} that implements the socially optimal research policy solves the equation[85]

$$\frac{c_e}{p_e} - c_i\Big(1-\beta(1-p_e)\Big) - \tau_{TC}\Big(\frac{c_e}{p_e} - c_i(1-\beta(1-p_e))\Big) = K - K\lambda\Big((1-\beta(1-p_e)\Big) \quad (90)$$

Under the simple research grant policy, that does not discriminate between young and established firms, the lowest matching rate for explorative projects τ'_{RG} that implements the socially optimal research policy is given by[86]

$$\frac{c_e}{p_e} - c_i\Big(1-\beta(1-p_e)\Big) - \tau'_{RG}\frac{c_e}{p_e} = K - K\lambda\Big((1-\beta(1-p_e)\Big) \quad (91)$$

Rearranging equation (13) from Proposition 5, it can be seen that the matching rate for the optimal research grant policy from Proposition 5 satisfies

$$\frac{c_e}{p_e} - c_i\Big(1-\beta(1-p_e)\Big) - \tau_{RG}\Big(\frac{c_e}{p_e} + \beta c_e\Big) = K - K\lambda\Big((1-\beta(1-p_e)\Big) \quad (92)$$

Hence, for $(\tau_{TC}, \tau'_{RG}, \tau_{RG})$ it holds true that

$$-\tau_{TC}\Big(\frac{c_e}{p_e} - c_i(1-\beta(1-p_e))\Big) = -\tau'_{RG}\frac{c_e}{p_e} = -\tau_{RG}\Big(\frac{c_e}{p_e} + \beta c_e\Big) =$$
$$K - K\lambda\Big((1-\beta(1-p_e)\Big) - \frac{c_e}{p_e} + c_i\Big(1-\beta(1-p_e)\Big) < 0$$

[85] (90) can be obtained by solving

$$\tilde{V}_Y = \tilde{V}^{obs}_{Estbl} = -c_e(1-\tau_{TC}) + p_e(K + \beta\tilde{V}^1_{Estbl}) + (1-p_e)\beta\tilde{V}_{Young}$$

and

$$\tilde{V}^1_{Estbl} = -c_e(1-\tau_{TC}) + p_e(K + \beta\tilde{V}^1_{Estbl}) + (1-p_e)(K\lambda - c_i(1-\tau_{TC}) + \beta\tilde{V}_{Young})$$

and plugging the resulting value functions into the reservation value condition for the established firm. The optimal tax credit satisfies the reservation value condition with equality.

[86] (91) can be obtained by solving

$$\tilde{V}_Y = \tilde{V}^{obs}_{Estbl} = -c_e(1-\tau'_{RG}) + p_e(K + \beta\tilde{V}^1_{Estbl}) + (1-p_e)\beta\tilde{V}_{Young}$$

and

$$\tilde{V}^1_{Estbl} = -c_e(1-\tau'_{RG}) + p_e(K + \beta\tilde{V}^1_{Estbl}) + (1-p_e)(K\lambda - c_i + \beta\tilde{V}_{Young})$$

and plugging the resulting value functions into the reservation value condition for the established firm. The optimal matching rate satisfies the reservation value condition with equality.

The last strict inequality follows from the assumption that exploration is not privately optimal for the established firm. The set of equalities above immediately implies the result. In addition, the set of equalities above can be used to compute the ratios of the matching rates.

Firm Sector and Firm Age Dummy Variables

In this section, I describe the firm sector and age dummy variables that are included as controls.

Firm Sector

The firm sector classification in this paper is based on the NACE 2 classification. The distinct sectors are listed further below. 80 percent of the firms are classified based on their NACE 2 code in AMADEUS. For firms without AMADEUS match, or that have an uninformative NACE 2 Code in AMADEUS (e.g. "7010 - Activities of Head Office"), I manually impute the sector based on information about the firm on the internet. First, I search the firm on www.firmenabc.at, www.unternehmen24.at and www.moneyhouse.at. If successful, I match the sector of operation mentioned on the site with the corresponding sector in my classification. Otherwise, I search the firm on www.google.com and try to find out about the sector of operation through the firm homepage or other news articles or business service sites that mention the firm. My sector classification is deliberately coarse to reduce the risk of wrong sector assignments during the imputation process. Still, for 4.5 percent of the funding applications, I was not able to determine the firm sector of operation. Such funding applications are assigned to the residual class "Unassigned".

The sector classification used in this paper, along with their corresponding NACE 2 codes and the share of funding applications in my baseline sample is as follows:

- Agriculture and Mining: NACE 2 Codes 01-09, 1.2%
- Manufacture of Food Products, Oil Products and Wood Products: NACE 2 Codes 10-19, 7 %
- Manufacture of Metal Products, Electronics and Chemical Products (without Pharmaceuticals): NACE 2 Codes 20,22-29, 41%
- Research Manufacturing: NACE 2 Code 7210, 1.9%
- Engineering Services ("Ingenieursbüro" in German): NACE 2 Code 7112, 5.3%
- Manufacture of Pharmaceutical Products: NACE 2 Code 21, 2%
- Research Pharmaceuticals: NACE 2 Code 7211, 2.8%

- Manufacture of Instruments, Sports Goods and other Equipment: NACE 2 Codes 30-34, 17%
- Power Supply and Construction: NACE 2 Codes 35-42, 3.1%
- Wholesale: NACE 2 Codes 43-46, 4.9%
- Consulting and Financial Services: NACE 2 Code 63, 4.4%
- Software: NACE 2 Code 62, 15.8%
- Rest: 2.6%
- Unassigned: 4.5%

The sizes of the sectors is quite unbalanced, with the smallest, Agriculture and Mining only having 33 observations. Collapsing the sectors with small shares into bigger sectors roughly following the Manufacturing/Engineering, Pharma, Software, other Services and Rest divide yields identical estimates. "Research Manufacturing" (NACE 2 Code 7210) and "Research Pharmaceuticals" (NACE 2 Code 7211) are NACE 2 Codes reserved for pure research firms, without production. Merging the sectors of pure research firms and manufacturing firms yields identical estimates. The NACE 2 classification also distinguishes pure research firms from firms that offer Engineering Services (NACE 2 Code 7112). Again, not making this distinction is completely inconsequential.

Firm Age

For 80 percent of the firms, the year of incorporation is assigned based on the respective entry in AMADEUS. For firms without AMADEUS match, or with an ineligigible entry (year of incorporation is equal 0 or 9999), I manually impute the year of incorporation based on information about the firm on the internet. First, I search the firm by name and address on www.firmenabc.at, www.unternehmen24.at and www.moneyhouse.at. All of these sources are based on the official Austrian firm register ("Firmenbuch")[87] and registers for one-man businesses that do not meet the sales threshold for the firm register. To the best of my knowledge, the sources track name changes and bankruptcies. As discussed in section 1.5.1, the age of the applicant firm (at the time of the funding application) is available for 2444

[87]Direct access to the register is restricted and it costs between 2 Euro and 10 Euro to obtain the entry of a single firm.

applications (93.3 percent of applications) in the baseline sample. The most plausible reason for why the age may be missing is that the firm never incorporated. For 10 percent of the applications in the sample, the firm age is negative, which means that the firm incorporated after the funding application. As mentioned in section 1.5.1, there is some noise in the measurement of firm age. If a firm changes its legal form, it reincorporates and resets the official year of incorporation. To the best of my knowledge, there is no way to distinguish genuinely new firms from firms that were formed as a result of a restructuring.

Firm age is defined as year of the funding application minus year of incorporation. Figure 19 and Figure 20 contain the distribution of firm age in the baseline sample (for applications with available firm age). The age groups, along with their relative shares, which are included as age fixed effects in all regressions, are:

- Younger than 2 years: 31.1%
- Between 3 and 6 years: 18.0%
- Between 7 and 15 years: 20.5%
- Older than 15 years: 30.3%
- Unknown age: 6.6%

Further refinements of the age groups yield identical results.

Patent measures based on citation data

In this section, I describe the patent measures used in section 1.5.2.

"Patent Unconventionality" was introduced by Berkes & Gaetani (2016). They compute the technical "relatedness" of IPC classes by calculating the frequency with which two IPC classes were cited by the same patent, using the entire population of patent applications to the USPTO. This symmetric measure for pairs of IPC classes called "c-score" is then used to calculate how unconventional the backward citation structure of any given patent is. I use Berkes & Gaetani's (2016) c-scores and calculate patent unconventionality for a large set of patents by European assignees. I include all patents in PATSTAT filed after 1980 by applicants with a country code from Germany, Italy, Norway, Sweden, Finland, Netherlands, Belgium, Denmark, Austria and Switzerland. Every patent is assigned the minimal c-score across all pairs of distinct IPC classes cited by the patent as the conventionality score, with no further normalization. Then, I rank all patents filed in the same IPC class and year by their conventionality scores and record whether a patent has a conventionality score below or above the median among all patents filed in the same IPC class and year. I classify patents with a below-median score as "Unconventional Patents" and patents with an above-median score as "Conventional Patents". I do not distinguish between National Patents and European Patents. In the sample, for firms of age greater than 5 years, the share of Conventional Patents in all patents filed between 1980 and the year of the funding application, is 47 percent. This share drops to 39 percent in the 4 years after (including the year of) the funding application.

"Patent quality" is measured by the number of forward citations that the patent received, in the vein of Trajtenberg (1990) and Hall et al. (2005). I utilize the same set of patents as above to rank all patents filed in the same year by the number of forward citations received, with no distinction between European Patents and National Patents. Patents with a number of forward citations higher than the median are classified as "High-quality Patents", whereas patents with a citation count below the median are classified as "Low-quality Patents".

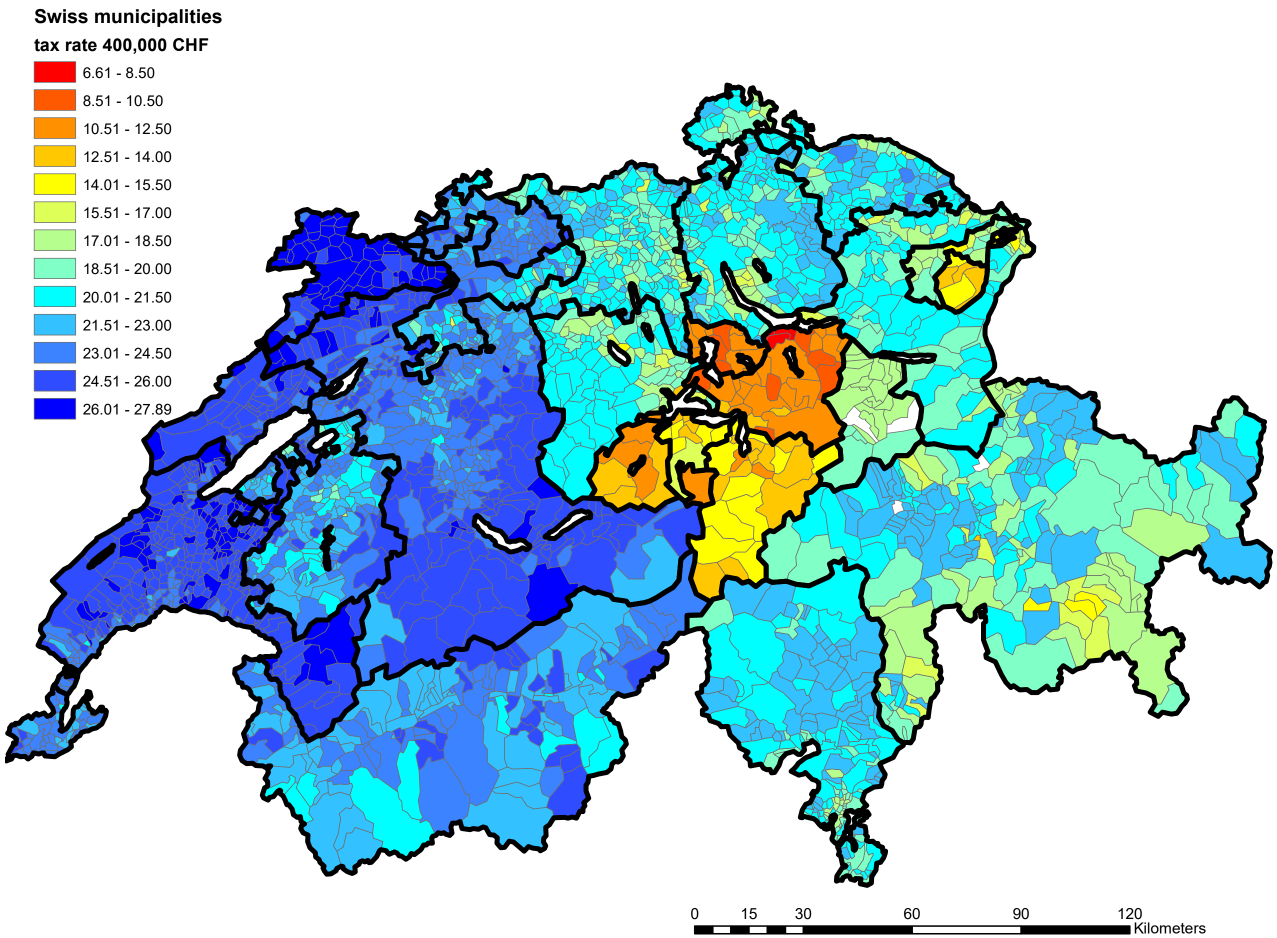
In analogy to the definition of European and National Patents (see section 1.3.3), Unconventional, Conventional, High-quality and Low-quality Patents are in fact DOCDB patent families that possibly comprise multiple patent applications. All citation counts are computed at the level of DOCDB patent families.

Chapter 2

Figure 25: Sample of posters from the party SVP ("Schweizerische Volkspartei") for elections and referenda related to immigration

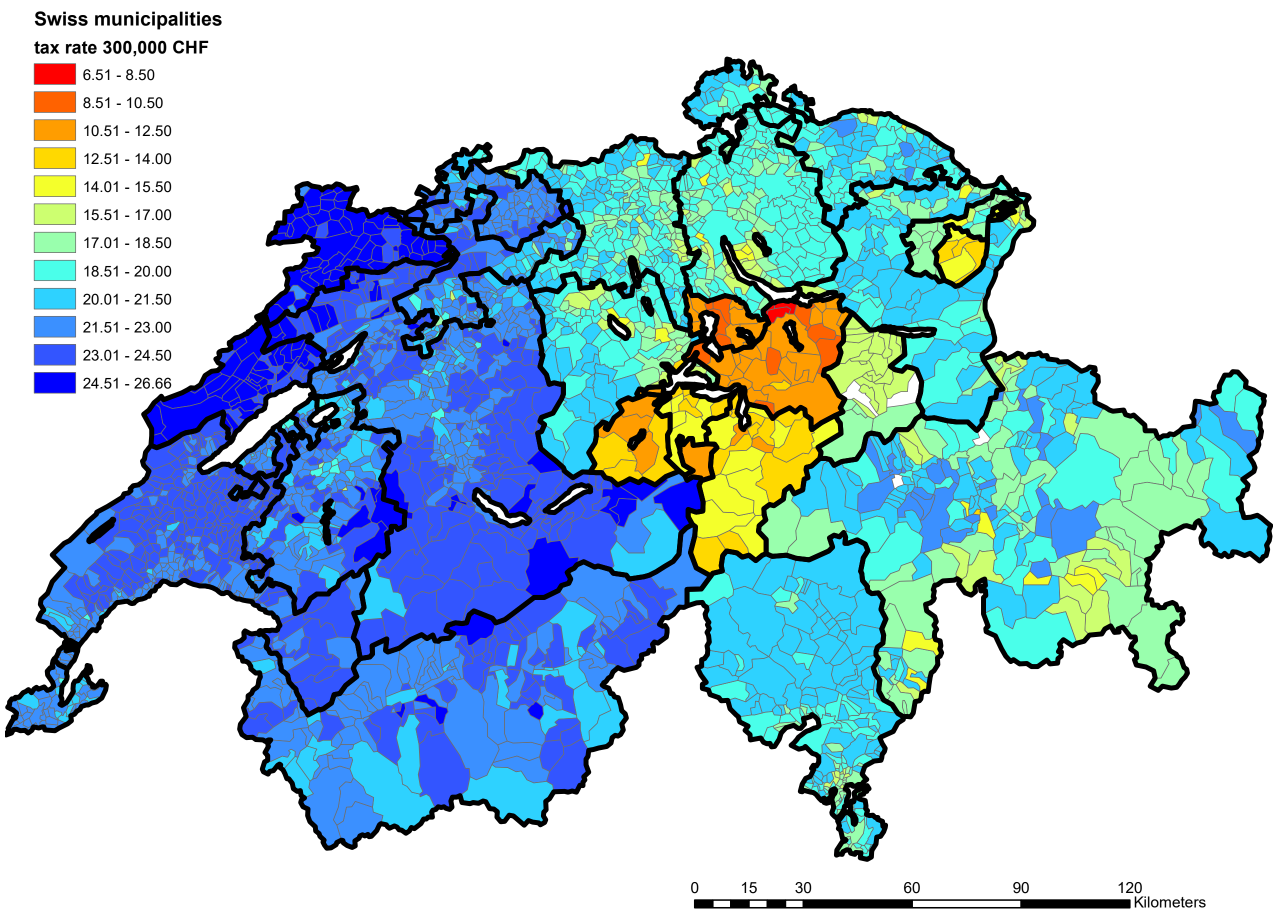
Note: First poster from the referendum "Gegen Masseneinwanderung" (in english: "Against mass immigration") in 2011, second poster from the referendum "Für die Ausschaffung krimineller Ausländer (Ausschaffungsinitiative)" (in english: "For the deportation of criminal foreigners (deportation initiative)" in 2007 and third poster in opposition to the referendum "Erleichterte Einbürgerung der dritten Generation" (in english: "Easening the requirements for the naturalization of third-generation immigrants").

Figure 26: Tax burden in percent of the gross income levied by the state and the municipality, excluding the federal income tax, for an unmarried individual with no children earning 400,000 CHF annually.



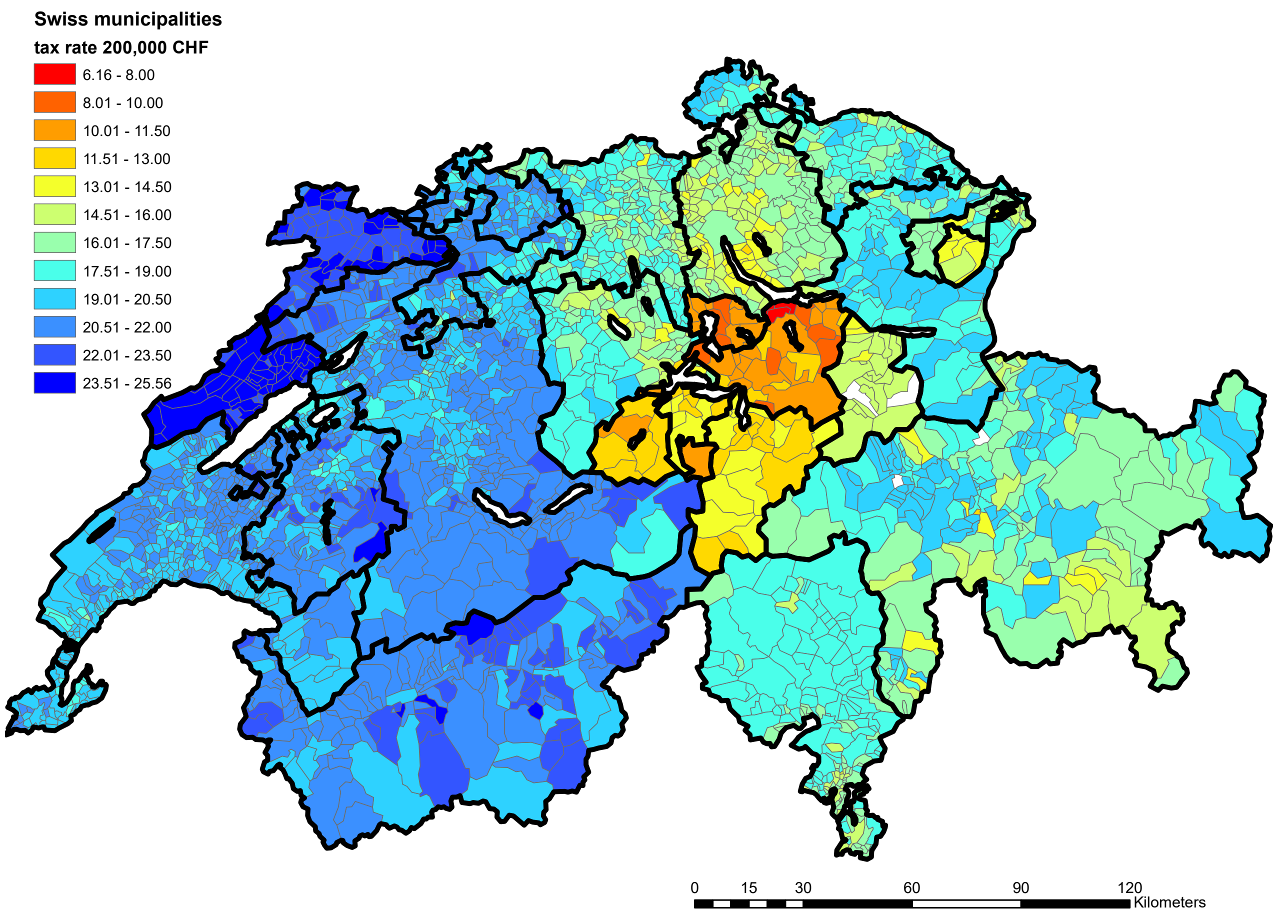
Note: Swiss municipalities as of 11-21-2010. Tax rates as of 01-01-2010. Tax rates missing for 4 out of 2584 municipalities (Matt, Luchsingen, Felsberg, Mundaun).

Figure 27: Tax burden in percent of the gross income levied by the state and the municipality, excluding the federal income tax, for an unmarried individual with no children earning 300,000 CHF annually.



Note: Swiss municipalities as of 11-21-2010. Tax rates as of 01-01-2010. Tax rates missing for 4 out of 2584 municipalities (Matt, Luchsingen, Felsberg, Mundaun).

Figure 28: Tax burden in percent of the gross income levied by the state and the municipality, excluding the federal income tax, for an unmarried individual with no children earning 200,000 CHF annually.



Note: Swiss municipalities as of 11-21-2010. Tax rates as of 01-01-2010. Tax rates missing for 4 out of 2584 municipalities (Matt, Luchsingen, Felsberg, Mundaun).

Table 48: Robustness Check: Results on income tax rates and the residential location choice of inventors for different assumed annual income levels

Dependent variable: choice of municipality of residence										
	100,000 CHF		150,000 CHF		200,000 CHF		300,000 CHF		400,000 CHF	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Subset of Inventors	All	All	All	All	All	All	All	All	All	All
log net-of-tax rate (100,000 CHF)	1.9034	1.8692								
log net-of-tax rate (150,000 CHF)			0.5082	1.9122						
log net-of-tax rate (200,000 CHF)					2.2248	2.7558				
log net-of-tax rate (300,000 CHF)							5.4918	4.6514		
log net-of-tax rate (400,000 CHF)									6.2552	4.8835
Linear controls for municipal amenities	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
Observations	12891	12891	12891	12891	12891	12891	12891	12891	12891	12891
avg. McFadden R-squared	0.0007	0.2390	0.0001	0.2390	0.0012	0.2399	0.0096	0.2424	0.0170	0.2438

Note: 95% confidence intervals reported, *** p<0.01, ** p<0.05, * p<0.1. Controls as in Table 11 column 2.

Table 49: Robustness check: Results on top income tax rates and other amenities in the residential location choice of inventors when I exclude small states, the state of Zürich or mountainous regions

Dependent variable: choice of municipality of residence						
	Excluding small states		Excluding Zürich		Excluding mountainous regions	
	(1)	(2)	(3)	(4)	(5)	(6)
log net-of-tax rate	2.8934	4.5384	2.176	3.6440	5.9239	3.8331
commuting distance (km)		−0.1615		−0.1874		−0.1853
population size (thousd.)		0.1339		0.1004		0.0995
lakefront (binary)		0.2211		0.4329		0.4270
distance national border (km)		−0.0189		−0.0219		−0.0199
Other controls: school grad. rate, crime rate, public transport usage	NO	YES	NO	YES	NO	YES
Observations	10353	10353	9721	9721	11969	11969
avg. McFadden R-squared	0.0029	0.2353	0.0019	0.2102	0.0190	0.2484

Note: 95% confidence intervals reported, *** p<0.01, ** p<0.05, * p<0.1. Controls as in Table 11 column 2. In columns (1) and (2), I exclude the states of Zug, Schaffhausen, Nidwalden, Obwalden, Appenzell Innerrhoden, Appenzell Ausserhoden and Uri and all municipalities that lie around the borders of those states. In columns (3) and (4), I exlude the state of Zürich and all municipalities that lie around the state border of Zürich. In column (5) and (6), I exclude all municipalities that are covered by 10 percent or more by rock.

Table 50: Robustness check: Results on top income tax rates and other amenities in the residential location choice of inventors for alternative specifications of the inventors' choice sets

Dependent variable: choice of municipality of residence						
	Municipalities within 7.5 km of distance to the state border		Municipalities within 20 km of distance to the state border (includes urban centers)		Municipalities that border the state	
	(1)	(2)	(3)	(4)	(5)	(6)
log net-of-tax rate	2.5344	3.6650	0.8431	4.1645	5.2893	4.6962
commuting distance (km)		−0.1596		−0.0875		−0.1618
population size (thousd.)		0.1188		0.0576		0.0615
lakefront (binary)		0.3194		0.3811		0.7007
distance national border (km)		−0.0399		−0.0276		−0.0292
Other controls: school grad. rate, crime rate, public transport usage	NO	YES	NO	YES	NO	YES
Observations	36608	36608	559003	5590038	11638	11638
avg. McFadden R-squared	0.0039	0.2992	0.0004	0.4567	0.0155	0.2766

Note: 95% confidence intervals reported, *** p<0.01, ** p<0.05, * p<0.1. Controls as in Table 11 column 2. In columns (1) and (2), I expand the inventors' choice sets to municipalities within 7.5 kilometers of distance to the inventor's municipality of residence (as opposed to 5 km). Municipalities that lie within 7.5 kilometers of distance to a municipality in a different state are included in the sample. In columns (3) and (4), I expand the inventors' hoice sets to municipalities within 20 kilometers of distance to the inventor's municipality of residence. Municipalities that lie within 20 kilometers of distance to a municipality in a different state are included. In particular, the urban centers of Zürich, Basel, Lausanne and Geneva are included in the sample. In columns (5) and (6), the inventors' choice sets are given by all municipalities that border the inventor's municipality of residence. Municipalities that border a municipality in a different state are included in the sample.

Description municipality data

In this section, I provide information on the municipal data in section 2.3.3. All data was obtained from the homepage of Swiss Statistical Office (www.bfs.admin.ch) in June 2017. The years in which the data was collected differs slightly by variable. Because the set of municipalities is changing over time, as villages are combined to form new municipalities or are reassigned to existing municipalities, I match all municipal data, whenever possible, to the set of municipalities as of 11-21-2010. If a municipality in 2010 corresponds to a combination of multiple municipalities in a different year, I impute the average across the respective municipalities. Conversely, if multiple municipalities in 2010 correspond to one municipality in a different year, I assign the value of the municipality to all corresponding municipalities in 2010. In more complicated cases, I refrain from imputations and discard the observation. However, overall, these are minor issues and concern only a handful of observations.

- School graduation rate: district-level share of 19-year olds who attained the degree "Gymnasiale Maturität", which is a requirement for university studies, in 2012.[88]

- Crime rate: number of violent and non-violent crimes per thousand population in the municipality in 2010.

- Municipality population size: population of the municipality in 2010.

[88] Districts are intermediate geographical units in between municipalities and states. In 2010 and 2012, there were 148 districts. The attainment rate ranges from 7.14 percent in rural Kulm (Aargau) to 42 percent in Lavaux-Oron (Waadt) near Lausanne. Schools are administered by states and students are allocated to public schools based on the municipality of residence. Parents may decide to opt-out of public schooling and instead enroll in a private school. Private schools are not subject to residence allocation rules and account for around 13 percent of the total number schools, evenly distributed across states. Due to the institutional complexity and because public authorities are actively trying to deter strategic sorting into schools, it is notoriously difficult to find data on school quality. Since access to universities is regulated federally, I expect the academic standards for attaining "Gymnasiale Maturität" to be similar across all states. The "Gymnasiale Maturität" is a rather exclusive degree whose attainment entails careful selection of schools from an early age on. In 2010 only 20 percent of 19 year olds in the general population attained "Gymnasiale Maturität".

- Distance to the national border: straight-line distance from the centroid of the municipality to the closest line segment of the national border.
- Public transport usage: district-level share of individuals who commuted by public transport in 2010.
- SVP vote share: share of votes in the municipality received by the SVP ("Schweizerische Volkspartei") in the federal elections of 2011.
- Language spoken in the municipality: share of population that named German or French as their primary language in 2000.