NBER WORKING PAPER SERIES

TAXATION AND THE INTERNATIONAL MOBILITY OF INVENTORS

Ufuk Akcigit
Salomé Baslandze
Stefanie Stantcheva

Working Paper 21024
http://www.nber.org/papers/w21024

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2015

We thank Jeremy Greenwood, Dirk Krueger, Jim Poterba, Emmanuel Saez, Ivan Werning, participants at the MIT Public Finance lunch and the Penn \Growth Reading Group" for useful comments and feedback. We thank Olof Ejermo, William Kerr, Francesco Lissoni, Ernest Migu elez, Valerio Sterzi, Myriam Mariani, Alfonso Gambardella for sharing their data. Akcigit gratefully acknowledges financial support from Ewing M. Kauffman Foundation and from the Alfred P. Sloan Foundation. We gratefully acknowledge financial support from the Washington Center for Equitable Growth. We thank Jacob R. Dorabialski for help with the analysis. We especially thank Nina Roussille for outstanding research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Taxation and the International Mobility of Inventors
Ufuk Akcigit, Salomé Baslandze, and Stefanie Stantcheva
NBER Working Paper No. 21024
March 2015, Revised October 2015
JEL No. F22,H21,H24,H31,J61,O33,O38

ABSTRACT

This paper studies the effect of top tax rates on inventors' international mobility since 1977. We put special emphasis on "superstar" inventors, those with the most abundant and most valuable patents. We use panel data on inventors from the United States and European Patent Offices to track inventors' locations over time and combine it with international effective top tax rate data. We construct a detailed set of proxies for inventors' counterfactual incomes in each possible destination country including, among others, measures of patent quality and technological fit with each potential destination. Exploiting the differential impact of changes in the top tax rate on inventors of different qualities, we find that superstar top 1% inventors are significantly affected by top tax rates when deciding where to locate. The elasticity to the net-of-tax rate of the number of domestic superstar inventors is relatively small, around 0.03, while the elasticity of the number of foreign superstar inventors is around 1. Inventors who work in multinational companies are more likely to take advantage of tax differentials. On the other hand, if the company of an inventor has a higher share of its research activity in a given country, the inventor is less sensitive to the tax rate in that country.

Ufuk Akcigit
Department of Economics
University of Chicago
1126 East 59th Street
Saieh Hall, Office 403
Chicago, IL 60637
and NBER
uakcigit@uchicago.edu

Salomé Baslandze
University of Pennsylvania
3718 Locust Walk
McNeil 160
Philadelphia, PA 19104
salomeb@sas.upenn.edu

Stefanie Stantcheva
Department of Economics
Littauer Center 232
Harvard University
Cambridge, MA 02138
and NBER
sstantcheva@fas.harvard.edu

1 Introduction

In 1876, Alexander Graham Bell invented the telephone, and created the Bell Telephone Company one year later. By 1886, more than 150,000 people in the United States owned telephones. In 1916, James L. Kraft patented a pasteurization technique for cheese and established his company, Kraft Foods Inc., that would grow into a conglomerate responsible for creating some of the United States' most popular food products and employing more than 100,000 people. In 1968, Ralph Baer created a TV game unit that allowed players to control on-screen action with paddle controls. Today, the video gaming industry is worth $66 billion. In the early 1970s, Michael Ter-Pogossian developed the positron emission tomography (PET) scanner, used today in countless medical examinations. In the mid-1970s, Samar Basu, through a series of patents, invented the technology that allowed the lithium ion batteries used in innumerable consumer products to be recharged multiple times. In 1981, Charles Simonyi started developing some of Microsoft Office's most profitable products. In addition to being very prolific inventors, these innovators had something else in common: they were all *immigrants*.

According to World Intellectual Property Organization data, inventors are highly mobile geographically with a migration rate around 8%.[1] But what determines their patterns of migration? In particular, how does tax policy affect migration? The fear of a "brain drain" and the exodus of economically valuable agents in response to higher taxation has led to a vivid public debate regarding the taxation of high income people. For instance, in response to the New York Times' (Feb, 2013) article entitled "The Myth of the Rich Who Flee From Taxes,"[2] following Gerard Depardieu's Russian exodus for tax purposes, Forbes issued an article entitled "Sorry New York Times, Tax Flight of the Rich Is Not a Myth."[3] In this paper, we study the effects of top income taxes on the international migration of inventors, who are key drivers of technological progress.

While an important issue, international migration responses to taxation have remained underexplored due to the lack of a large-scale international panel dataset. One important exception is the study of the migration responses of football players, a set of economic agents very different from inventors, by Kleven, Landais, and Saez (2013). In our analysis, we use a unique international panel data on all inventors from the U.S. and European patent offices in an unusual way, namely to track the international location of inventors since the 1970s. The benchmark data is a panel data from the Disambiguated Inventor Data by Lai *et al.* (2012), based on inventors who patent with the United States Patent Office (USPTO). The focus is on the 8 OECD countries that represent the bulk of USPTO patents for the period 1977-2000: Canada, France, Germany, Great Britain, Italy, Japan, Switzerland, and the United States. The new disambiguated European inventor data is from Coffano and Tarasconi (2014). The U.S. and European patent offices together account for a very large fraction of worldwide patents, so that our sample contains most of the universe of

[1] For a recent study using this data, see Miguelez and Fink (2013). High skilled workers are in general more mobile that low skilled workers, with inventors being among most high-skilled immigrants.

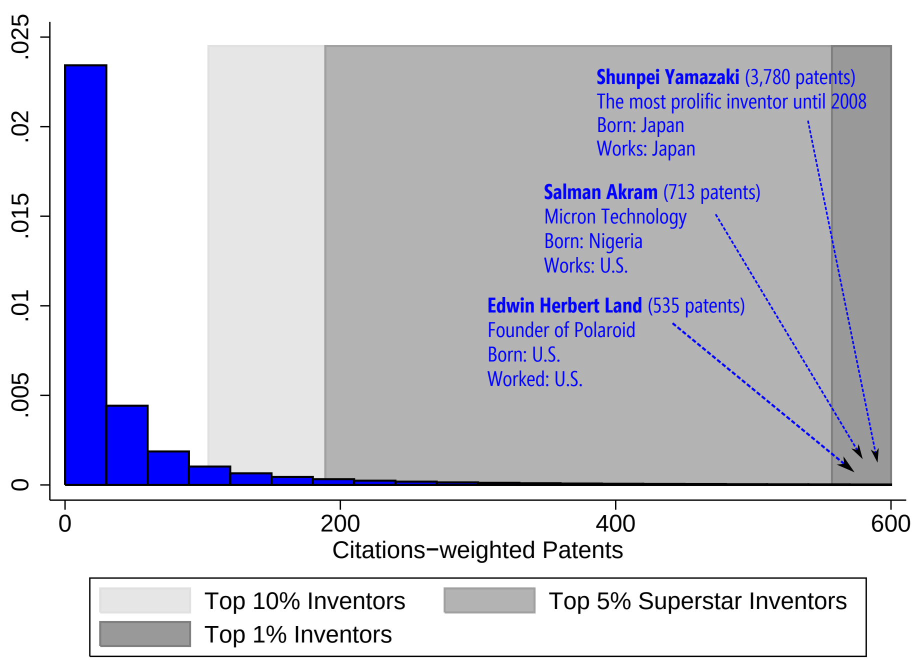
[2] Article by James B. Stewart, published February 15, 2013.

[3] Article by Paul Roderick Gregory, published February 17, 2013.

inventors who patent. We combine this inventor data with international effective marginal top tax rate data from Piketty, Saez, and Stantcheva (2014).

Particularly interesting are "superstar" inventors, namely those with the most abundant and most valuable innovations. The distribution of inventor quality, as captured by the citations that an inventor's patents receive– an often-used measure of the economic value of patents – is highly skewed:[4] As can be seen in Figure 1, while the median and mean inventors have, respectively, 11 and 42 lifetime citations, the average top 1% superstar inventor has 1019 citations.

Figure 1: Distribution of citations-weighted patents



Notes: This figure plots the distribution of citations-weighted patents (see formula (1)) across inventors in the U.S. Patent Office data 1977-2000 from 8 countries: Canada, Great Britain, Germany, France, Italy, Japan, Switzerland, and the United States. For the sake of visual clarity, the x-axis is truncated at 600 citations. For a detailed description of the data, see Section 2.3.

There are three major challenges that arise when studying the migration responses to taxation. These challenges come in addition to the aforementioned usual obstacle of lack of international panel data – a hurdle we are able to resolve in this paper thanks to the use of the patent data. First, conceptually, migration decisions should depend on the *counterfactual* income that an agent expects to receive in each potential location, which is naturally not observed. Second, migration decisions should depend on the counterfactual average tax rate that the agent expects to pay in each potential location, which itself depends on the (unknown) counterfactual income earned. Finally, average tax rates may be endogenous to other factors in a country and in a given year.

Our strategy to study mobility proceeds in three steps in order to address these three chal-

[4]See Section 2.3 for references to the literature studying patent citations.

lenges, making use of some unique features of the patent data. First, we construct detailed proxies for inventors' counterfactual earnings in each potential destination country, which include, among others, measures of inventors' qualities. Indeed, the patent data provides us with a rare opportunity to directly measure inventor productivity and quality. Our benchmark measure of quality is citations-adjusted patents, but we also consider the number of patents, average citations per patent, the maximum citations per patent, as well as measures of the breadth of patents and breadth of impact of an inventor.[5] We focus on inventors who are employees of companies and, hence, not the owners of their patents.

The quality measures based on citations, are important determinants of inventors' incomes, directly and indirectly. Directly, more citations mean a higher economic value of the patent (Trajtenberg, 1990) and companies would typically reward inventors through bonuses or "fair-share" agreements. Indirectly, employers should try to promote and retain their star employees, as characterized by high patent qualities (see also the papers reviewed in Section 2.4 which document the link between patent citations and income).

In addition to inventor quality, we also proxy for counterfactual earnings through a variety of variables, such as inventors' technological field, their technological fit with each potential destination country, their tenure, and ability-technological field-country specific trends. We also exploit information about what type of institution the inventor works for, notably whether he is employed by a multinational, and the share of the innovative activity of his company performed in each potential destination country.

Based on our benchmark quality measure (and for each of the other measures), we construct a corresponding quality distribution, conditional on region of origin and year, and determine the rank of each inventor in this distribution.[6] We define superstar inventors as those in the top 1% of the quality distribution, and similarly construct the top 1-5%, the top 5-10%, and subsequent quality brackets.

Second, regarding the counterfactual average tax rate, we focus on the elasticity of migration to the top marginal income tax rate. The resulting elasticity is a reduced form parameter and our estimates are not necessarily interpretable as the "elasticities of migration to net-of-tax income," for several reasons. First, top marginal tax rates are not equal to effective top average tax rates because of the nonlinearity of the tax system. This should be a minimal concern in the case of very high earners who are well above the top tax bracket. Second, there are other potentially relevant taxes for inventors, such as capital or corporate taxes.[7] To minimize the latter concern,

[5] See Section 2.3 for a detailed description of how we construct quality measures for inventors and for alternative definitions of "superstars."

[6] The three benchmark regions used to construct the per region and year quality distributions are based on comparable patenting intensities. They are i) the U.S., ii) European countries and Canada, and iii) Japan, but we also use per-country rankings for robustness checks (in Appendix Table A16).

[7] Note however that we are not trying to understand the effects of the taxation of *direct* patent income (such as royalties or corporate income). We are instead interested in the taxation of salary and bonus earnings, for which personal income taxes are the main consideration. Appendix Table A14 shows that the effect of capital gains taxes on migration is insignificant once income taxes are controlled for.

we restrict our benchmark sample to those inventors who are employees of companies and who are not the assignees of the patents – this ensures that personal ordinary income tax should be the main tax affecting decisions.[8] Furthermore, firms' decisions and institutional factors (such as the visa system) influence the response to taxes. Incidentally, the estimated elasticity to the top tax rate may also be interesting *per se* as it captures the effects of a "success" tax.

Third, turning to the identification, the simplest strategy is to exploit variation in top marginal tax rates across time and countries, controlling for country fixed effects, year fixed effects, and country specific linear trends to filter out longer-term country-specific evolutions in innovation, incomes, and migration. We do this analysis merely as a first pass as there may be other factors correlated with top marginal tax rates which vary at the same time in a given country. Instead, our preferred identification filters out all country-year level variation and exploits the differential impact of changes in the top marginal tax rate on top superstar inventors and slightly lower quality inventors within a given country and year. Top 1% inventors and those of slightly lower qualities are similar enough to be subject to the same country-year level policies and economic effects. However, only inventors who are actually in the top bracket are directly affected by the changes in top tax rates. The evidence presented suggests that top 1% superstar inventors are well in the top tax bracket. While the top 1-5% are still very likely to be in the top bracket, the likelihood of being in the top tax bracket declines sharply as we move down through the top 5-10% to the top 10-25% or below top 25% inventors.[9] Hence, the lower quality, top 5-10%, top 10-25%, and below top 25% groups serve as control groups for the top 1% group.[10]

We perform the analysis at three, successively more detailed levels. We start by documenting some stylized macroeconomic facts about the relation between mobility and top tax rates. Since the 1970s there is a negative correlation between the top tax rate and the share of top quality foreign inventors who locate in a country, as well as the share of top quality domestic inventors who remain in their home country.

Delving deeper into three specific case studies, we exploit the quasi-experimental variation provided by migration or tax reforms. First, we consider the special case of Russian inventors–

[8] We also check that our results hold for the full sample which includes non-employees in Appendix Table A16.

[9] As explained in Section 5.1, "below top 25%" refers here to a relatively high quality group of inventors who at some point have been or will be in the top 25%, but are not currently in the top 25%.

[10] Conceptually, this main identification strategy is akin to a "fuzzy regression discontinuity design," where the quality ranking generates a fuzzy threshold above which inventors are "treated" by the top tax rate changes. The design is fuzzy because we do not know counterfactual income and use quality rankings instead. This is similar in spirit to papers using the differential impact of tax changes across income tax brackets (Eissa, 1995) or across households with different non-labor income and family size (Eissa and Hoynes, 2006) rather than quality brackets like we do. It is most closely related to the differential propensities of being "treated," as in Hoynes and Schanzenbach (2009) or Hoynes and Schanzenbach (2012) who multiply Food Stamp program introduction by propensity to participate (proxied by group-level food stamp participation): in our case, we interact the top tax rate with the inventor's quality ranking (which proxies the propensity of being in the top bracket). When choosing the control group for top 1% inventors, there is a trade-off between maximizing comparability (i.e., choosing a group very close in quality e.g., the top 5-10%) and minimizing the control group's propensity of being treated (i.e., choosing a group further down the quality distribution, e.g., the below top 25%). We therefore report the estimated coefficients for all inventor quality groups and the implied elasticities relative to three different choices of control group.

a group whose migration opportunities were severely restricted before the collapse of the Soviet Union. We show that, in accordance with our main identification strategy, top quality Russian inventors, relative to lower quality Russian inventors, tended to migrate to lower top tax countries after the Soviet Union collapsed. We then exploit two large tax reforms, namely the U.S. Tax reform Act of 1986 and Denmark's "Researchers' Tax" reform of 1992.

Finally, we estimate a full-fledged multinomial location choice model to obtain the micro elasticities to top tax rates for foreign and domestic inventors. We use the two identification strategies previously described exploiting, first, country-by-year variation in top tax rates and, second, the differential effective impact of top tax rates on inventors of different qualities.

We find that the superstar top 1% inventors are significantly affected by top tax rates when choosing where to locate. As explained in the text, we provide the effects relative to several possible control groups for the top 1% inventors. At the lower end, using the top 5-10% as a control group yields an elasticity of the number of domestic superstar top 1% inventors to the net-of-tax rate of 0.02, which, as explained below, may be a lower bound. Using the top 10-25% or the below top 25% as control groups yields elasticities of, respectively, 0.02 and 0.03. On the other hand, the elasticity of the number of foreign top 1% superstar inventors to the net-of-tax rate is much larger, with corresponding values of 0.63, 0.84, and 1. To put these numbers into perspective, the elasticity of the number of domestic football players in Kleven, Landais, and Saez (2013) to the net-of-tax rate ranges from 0.07 to 0.16 depending on the specification, and the elasticity of the number of foreign players ranges from 0.6 to 1.3. We also find some evidence consistent with sorting by ability and general equilibrium wage effects.

The type and structure of companies seems important for the migration decision: Inventors who have worked for multinationals in the previous period are more likely to take advantage of tax differentials. On the contrary, they are much less sensitive to the tax rate in a given country if their company has a significant share of its innovative activity in that country.

We then perform extensive robustness checks on these benchmark results. First, we consider five different measures of inventor quality, both static and dynamic, and alternative proxies for earnings. Second, we contrast short-term and long-term mobility in response to taxation. Third, we address potential selection based on patenting behavior by estimating a Heckman selection model that exploits a 1994 reform effectively expanding patent protection in the U.S. Finally, we reproduce the analysis on the disambiguated inventor data from the European Patent Office and find very similar elasticities of the number of domestic and foreign top 1% superstar inventors to the net-of-tax rate of, respectively, 0.02 and 0.76.

Related Literature: That high skilled migration and its drivers are important considerations has been highlighted in the literature (see the paper by Kerr (2013) and the extensive references therein). High-skilled immigrants account for roughly 25% of U.S. workers in innovation and entrepreneurship and contribute disproportionately to patents or start-ups. Immigrants account for a majority of the net increase in the U.S. STEM workforce since 1995. There seems to be a positive selection among immigrants, as documented by Grogger and Hanson (2011) for foreign Ph.D. students who stay in the U.S. and by Chiquiar and Hanson (2005) for Mexican immigrant workers.

In turn, not just the receiving, but also the sending country is affected. It could either be suffering from a "brain drain," by losing highly skilled people or gaining from a "brain gain," if people invest in human capital more with the prospect of immigration or from remittances. Indeed, oversea diasporas have been studied as important determinants of knowledge flows (Foley and Kerr, 2013). Inventor migration and the formation of geographical knowledge clusters and their spillovers have also received attention (Miguelez and Moreno (2014), Miguelez (2013), Breschi, Lissoni, and Tarasconi (2014) and the references therein).

Our paper adds to a recent literature that studies the international migration of people in response to taxation. Most closely related are the papers by Kleven, Landais, Saez, and Schultz (2014) and Kleven, Landais, and Saez (2013). Kleven, Landais, Saez, and Schultz (2014) find very high elasticities of the number of high income foreigners in Denmark using a preferential tax scheme on high-earning foreigners implemented by Denmark in 1992 that reduced top tax rates for 3 years.[11] Kleven, Landais, and Saez (2013) study the migration of football players across European clubs. While we find somewhat lower elasticities for domestic inventors, Kleven, Landais, and Saez (2013) themselves conjectured that football players might be substantially more mobile than other high-skilled workers, because they earn most of their lifetime income over a short period and their profession involves little country-specific capital. In addition, their sample exclusively considers migration across European countries, while we also include the United States, Canada, and Japan. Expanding the study to other continents might, one would expect, reduce the tax elasticities of migration. Other related papers consider migration within countries. Bakija and Slemrod (2004) use Federal Estate Tax returns to show that the effect of higher state taxes on the migration of wealthy individuals across states in the U.S. is very small.[12] Moretti and Wilson (2014) consider aggregate state level effects of adopting subsidies for biotech employers –such as increase in R&D tax incentives– on the inflows of star scientists. Very complementary to our paper is the recent creative paper by Moretti and Wilson (2015) which considers the effects of state taxes on the migration of star scientists across U.S. states and also finds highly significant effects of taxes on migration.

Estimating the elasticity of migration to the tax rate is also important for the theoretical literature of optimal taxation with migration, as it enters the optimal tax formulas (Mirrlees, 1982; Wilson, 1980, 1982).[13] Our paper can shed some light on how big the proposed modifications to standard optimal tax formulas have to be empirically in order to account for migration.[14]

[11]By contrast, Young and Varner (2011) study the effects of a change in the millionaire tax rate in New Jersey on migration and find small elasticities.

[12]Liebig et al. (2007) study mobility within Switzerland, across cantons and find small sensitivities to tax rates.

[13]See also the more recent papers by Simula and Trannoy (2010) and especially Lehmann et al. (2014) who consider optimal nonlinear income taxation in the presence of migration.

[14]The same applies for the macro structural literature that includes migration channels and needs to calibrate the

The rest of the paper is organized as follows. Section 2 presents the setting and data, as well as a simple model of inventor migration. Section 3 shows some stylized macro facts on the relation between top tax rates and superstar inventors' migration. Section 4 presents quasi-experimental evidence based on three country case studies. Section 5 describes the multinomial location model estimation and the main results. Section 6 contains several robustness checks and extensions. Section 7 repeats the analysis on inventors who patent with the European Patent Office. Section 8 concludes.

2 Setting, Data, and Strategy

This section provides some background information on inventors, patents, and our various datasets. We explain how we use the unique features of patent data for the study of the effects of taxation on inventor mobility with the help of a simple location choice model.

2.1 Inventors and patents: Background

Inventors are the authors of innovations. They can be employees of companies, work for research institutions, or be self-employed "garage inventors." Patents protect the intellectual property of the innovation. They are legally granted to an assignee. The assignee is either an individual (possibly, but not necessarily, one of the inventors on a given patent), a national, local, or state government, an institute, a hospital or medical institute, or a university.

Inventors seem to be more mobile than the general population, which is consistent with a positively documented relation between skill and mobility. As summarized in Miguelez and Fink (2013), the global migration rate in 2000 for the population above 25 years old was 1.8%, ranging from 1.1% for the unskilled to 5.4% for those with tertiary education. In our benchmark sample, 2.3% of inventors move at least once over their lifetime in the sample (an average of 12 years) and 4.6% of the superstar inventors move over the same duration.

2.2 A Simple Model of Inventor Migration

To motivate our empirical analysis, consider the following very simple model of inventor migration. There are C countries, labeled as $c \in \{1, ..., C\}$. The wage of inventor i in country c' at time t is denoted by $w^i_{c't}$. Let $\tilde{w}^i_{c't}$ denote the corresponding marginal product. Index i allows the marginal product to depend on several inventor characteristics, such as his technological class, his age, as well as characteristics of the firm or employer he works for. If the international labor market is perfectly competitive, each inventor is paid his marginal product, so that $\tilde{w}^i_{c't} = w^i_{c't}$, $\forall c', t, i$.

Suppose that in country c, an inventor with home country h^i has to pay a tax rate τ_{ch^it} on his total income at time t. The tax rate is allowed to depend on the country of origin because foreigners

migration elasticities (see Cosar, Guner, and Tybout (2010)). The migration channel could also further bolster recent findings that the room for higher tax revenue through more progressive taxes is limited in Guner *et al.* (2014).

can sometimes face different tax regimes in different countries. For instance, U.S. citizens are taxed on their worldwide income.

In addition to the income earned, there is also a net utility benefit denoted μ_{ct}^{i} from locating in country c at time t for inventor i. This benefit is person-specific and can include, among others, a home bias, technological strengths and characteristics of the country, language differences, or distance to the home country. It can also capture country-specific characteristics of the company the inventor works for, such as the share of innovative activities the company performs in country c at any given time.

Total utility from choosing country c at time t for inventor i is given by:

$$U_{ct}^{i} = u\left(w_{ct}^{i}\left(1-\tau_{ch^{i}t}\right)+\mu_{ct}^{i}\right)$$

If to a first order there are no adjustment costs of moving, the decision of where to locate every period is history-independent. Hence, country c will be chosen in period t if and only if $u\left(w_{ct}^{i}\left(1-\tau_{ch^{i}t}\right)+\mu_{ct}^{i}\right) \in \operatorname{argmax}_{c'}\left\{u\left(w_{c't}^{i}\left(1-\tau_{c'h^{i}t}\right)+\mu_{c't}^{i}\right)\right\}$.

This dispenses us from having to make potentially unrealistic assumptions about the expectations of future tax rates, on which there is little empirical evidence and which could be country-specific. Note that for any two countries, none of which is the home country of the inventor, the utility from living in one of these countries does not depend on the other country – this is the sense in which there are no adjustment costs. However, the utility from living in any country does depend on the home country, for instance, through a home bias, geographical distance to or language differences with the home country (which are the factors we control for in the estimation). In that sense, the home country plays a special role because it enters as a time-invariant individual characteristic through the index i.

The model highlights that, conceptually, it is the average tax that should matter for the supply of inventors in a country. The probability that inventor i locates in country c will normally depend on the full vector of tax rates in all countries, $(\tau_{1h^{i}t}, ..., \tau_{ch^{i}t}, ..., \tau_{Ch^{i}t})$. However, for the empirical analysis, we make the assumption that the tax rate of any other country only has a negligible impact on the supply of inventors in country c. This is a good approximation if there are many possible origin and destination countries and each is relatively small. Hence, to a first order, the probability that inventors from country k locate in country c will depend only on τ_{ckt} and the relationship should be negative.

Note that the personal location preferences captured by μ could be so strong that they completely dominate any considerations of tax differences. As a result, strong location preferences unrelated to net income will reduce the observed sensitivity to tax rates. In the empirical analysis, we will explore such factors related to the type of job and company the inventor works in, in addition to the standard factors such as home bias, language differences, and geographical distances.

If the labor market for inventors is not perfectly flexible, and employers instead have a rigid demand for inventors, the wage need not be equal to the marginal product and may be a function of the tax system. This type of general equilibrium effects can be one of the potential reasons for endogeneity of top tax rates. We return to this issue in Sections 5.1 and 5.2.

2.3 The Inventor Data

Our main data source is the Disambiguated Inventor Data (hereafter, DID) by Lai *et al.* (2012), which identifies unique inventors in the U.S. Patent Office (USPTO) data. The USPTO data contains 4.2 million granted patents and 3.1 million inventors for the period 1975-2010, which represents 18% of worldwide direct patent filings and 26% of all patents.[15]

We limit the sample to the 8 major countries which account for 89% of all patents granted by the USPTO.[16] The U.S. accounts for 55% of the USPTO patents, Canada for 2.3%, Great Britain for 3%, Germany for 7.6%, Italy for 1.2%, Japan for 19.6%, France for 2.9%, and Switzerland for 1.3%. These representation differences reflect the different propensities that countries have for filing a patent with the USPTO: 58% of U.S. patent filings, 48% of Canadian filings, 19% of British filings, 16% of German filings, 20% of Italian filings, 13% of Japanese filings, 17% of French filings, and 12% of Swiss filings are filed with the USPTO.[17]

But filing propensities and representation in the patent data are not necessarily correlated with migration propensities. Indeed, while it is true that the largest migration corridors are the Great Britain-U.S. and Canada-U.S. corridors, other migration corridors such as the Japan-U.S. and Switzerland-U.S. ones are very small, although these countries have a high propensity of filing patents in the U.S..

The DID contains inventors' disambiguated names and residential address, which allows us to track the location of the inventor over time. Each of the inventor's patents has a patent number, application year, grant year, and assignee name. In addition, to get information on each patent's characteristics, such as citations or technological class, we merge the inventor data to the NBER patent data. Because of the truncation issue for patents' citations (more recent patents mechanically have less time to accumulate citations), we limit the sample to the years 1977-2000. Table 1 provides some summary statistics for the inventor data.

Only very recently has an effort of disambiguation similar to that done for the DID been undertaken for the European patent data (Coffano and Tarasconi (2014), Breschi, Lissoni, and Tarasconi (2014)). Despite the fact that there has been less work on this data and that the disambiguation algorithm is still less-well established, it offers some good advantages. The biggest of these is that the U.S. is less represented and European countries are more represented. We analyze inventor mobility using this data as well in Section 7.

For the macro stylized facts in Section 3 and the case studies in Section 4, we also use a third

[15] Patents can be filed with several offices and a direct patent filing is one of these recorded filings, while a patent is a given, unique invention.

[16] Appendix Table A16 shows that our results persist if we do not limit the sample to those countries.

[17] Source: Authors' calculations based on WIPO data available at: `http://ipstats.wipo.int/ipstatv2/index.htm?tab=patent`.

Table 1: Summary Statistics

Variable	Average
Patents of Superstar (Top 1%) Inventors	54
Patents of Superstar (Top 5%) Inventors	29.3
Patents of Non-superstar (Below Top 5%) Inventors	3.5
Average patents per year while in sample	1.5
Max citations per patent of Superstar (Top 1%) Inventors	147
Max citations per patent of Superstar (Top 5%) Inventors	100
Max citations per patent of Non-superstar (Below Top 5%) Inventors	24
Number of Patents (per country per year)	12,454
Number of Inventors (per country per year)	17,275
Number of Co-Inventors (per patent)	1.2
Number of immigrants (per country per year)	102
Number of immigrants per year to the U.S.	439
Number of immigrants per year to CA	71.2
Number of immigrants per year to CH	50.3
Number of immigrants per year to DE	78.6
Number of immigrants per year to FR	37.9
Number of immigrants per year to GB	87.3
Number of immigrants per year to IT	12.12
Number of immigrants per year to JP	34.5
% Superstar (Top 1%) Inventors who move over life in sample	4.6%
% Superstar (Top 5%) Inventors who move over life in sample	3.6%
% Non-superstar (Below 5%) Inventors who move over life in sample	0.7%
Average duration of stay in years conditional on move (benchmark sample)	5.3
% of inventors who are employees	83.2%
% of employees who work for multinationals	75%
Average years between first and last patent (benchmark sample)	12

Notes: Summary statistics are based on inventor and patents data set described in Section 2.3 for the period 1977-2000. The data includes inventors in 8 countries: Canada, France, Germany, Great Britain, Italy, Japan, Switzerland, and the United States. The sample contains 4,154,792 observations with 1,868,967 unique inventors. The benchmark estimation sample contains all inventors who have ever been in the top 25%.

alternative data source on inventors' locations. This dataset is described in detail by Miguelez and Fink (2013) and extracted from patent applications that are filed under the Patent Cooperation Treaty (PCT), a treaty administered by the World Intellectual Property Organization (WIPO) that offers some advantages for seeking international patent protection. The PCT data contains 54% of all international patent applications, but accounts for only 8% of worldwide patent filings. However, there has been no inventor name disambiguation and it is not yet possible to track inventors by name over time in a panel. As a result, we cannot construct dynamic quality measures for inventors as will be described below. The migration counts could be somewhat biased as, for instance, every inventor could be counted several times as a migrant. However, the great advantage of this dataset is that it contains nationality information. In addition, many smaller countries which are essentially non-existent in the DID are better represented in the PCT data, which allows us to provide results for a larger set of countries than our benchmark 8 countries. Thus, this data serves as a robustness check on the results using our benchmark data, while also providing an independent and different angle for the analysis.

Constructing quality measures for inventors: Citations received have traditionally been used as measures of the economic and technological significance of a patent (see Pakes (1986), Pakes and Schankerman (1986), Trajtenberg (1990), Harhoff et al. (1999), Hall, Jaffe, and Trajtenberg (2001), Bessen (2008), Kogan et al. (2012), Moser et al. (2012), Abrams, Akcigit, and Popadak (2013)). We construct four different dynamic measures of the inventor's quality, which place different importance on the quantity versus value of an inventor's patents. Let p_{ij} be the number of truncation-adjusted forward citations received by patent j of inventor i. Note that this does not depend on time t, as it counts all the forward citations that patent will ever receive, not the citations received until time t. The truncation adjustment, which takes into account the fact that more recent patents have less time to accumulate citations, is described in Hall, Jaffe, and Trajtenberg (2001). Let P^i_{t-1} be the set of patents of inventor i by the end of period $t-1$. Our benchmark measure is the lagged citations-weighted dynamic patent stock of the inventor. Formally, we denote this measure by $q1^i_t$ and it is equal to:

$$q1^i_t = \sum_{j \in P^i_{t-1}} p_{ij} \tag{1}$$

Measure $q1$ takes into account both the quantity and the quality of an inventor's patents, focusing on citations accumulated as a measure of one's influence. Our second measure, denoted by $q2^i_t$ is the lagged patent count of the inventor namely:

$$q2^i_t = |P^i_{t-1}| \tag{2}$$

where $|P^i_{t-1}|$ is the cardinality of the set P^i_{t-1}. This measure ignores the quality of patents and purely focuses on their quantity and is hence not our preferred measure.[18] The third measure, $q3^i_t$, is the lagged mean number of citations per patent:

$$q3^i_t = \frac{\sum_{j \in P^i_{t-1}} p_{ij}}{|P^i_{t-1}|} \tag{3}$$

which measures the average quality of an inventor's inventions to date. The fourth measure, $q4^i_t$, is the max number of citations ever received on a patent by inventor i:

$$q4^i_t = \max_{j \in P^i_{t-1}} p_{ij} \tag{4}$$

[18]Many patents have no real economic value and are never cited by any subsequent patent (see for instance Abrams, Akcigit, and Popadak (2013)).

which captures the best an inventor has ever achieved and whether he ever had a "home-run" invention. The additional results for our non-benchmark measures $q2$, $q3$, and $q4$ are in Section 6.

Based on these quality measures, we can define a ranking for inventors and, in particular, identify "superstar" inventors. We could in principle use a worldwide ranking of inventors. However, the patenting intensity is quite different for different countries and thus the quality measures are not necessarily directly comparable at a global level. This is why we group our 8 countries into 3 regions based on comparable patenting intensity: 1) the U.S., 2) Japan, 3) European countries and Canada. The U.S. and Japan stand out as the biggest patenting countries with 55% and 26% of all granted patents in the sample period (1977-2000).[19] We assign each inventor to a region based on whether his "home" country is in that region. Since we do not observe actual nationality in this data, we call home country the country in which the inventor is first observed in our sample. We define superstars at time t as those in top 1% of the regional quality distribution at time t. The top 5%, top 10%, and top 25% are calculated in a similar way. These are dynamic measures in the sense that they can change over an inventor's life, depending on where he falls in the regional distribution of quality in any given year.[20] Henceforth, we use the notation "top 1-5%" to denote inventors who are in the top 5% excluding the top 1%, and, similarly the top 5-10% and top 10-25% to respectively denote inventors in the top 10% excluding the top 5%, and in the top 25%, excluding the top 10%. Whenever we write "below top 25%" we refer to the inventors who have been or will be in the top 25% during their lifetime, but are not currently in the top 25%.

2.4 Making use of the Inventor and Top Tax Rate Data

There are three main challenges when studying migration: Proxying for the counterfactual earnings in each potential destination, measuring the counterfactual tax rate, and finding exogenous variation in the tax rate. Accordingly, we take a three step approach.

Step 1: Proxying for counterfactual earnings. Our model highlighted that migration decisions should conceptually depend on the income an inventor expects to earn in each potential destination, which is a counterfactual, unobservable variable.[21]

Fortunately, the patent data gives us a rich set of measures that can proxy for an inventor's counterfactual earnings. Among them are the previously described quality measures, $q1 - q4$, in formulas (1)-(4). Patent quality to date is a composite, dynamic statistic that takes into account an inventor's past achievements. In that sense, it is a measure of inventor ability or earnings potential, and a reflection of the inventor's "resume."

Patent quality and citations should increase inventors' incomes in both a direct and an indirect way. First, there are direct rewards and bonuses for specific innovations, driven potentially by legal

[19] The other countries each account from 1.16% to 8.85% of patents. Furthermore, the mean number of patents per inventor in the U.S., Japan and the rest of the countries is, respectively, 3.95, 4.7 and 3.3.

[20] As a check, we also defined the reference distribution and ranking separately for each country, instead of by region, and the results, available on demand, were virtually unchanged.

[21] Grogger and Hanson (2015) show that high ability immigrants are more likely to stay where there is a high premium for high skill workers. Hence, proxying well for the counterfactual wage is crucial.

or contractual arrangements, such as "fair share" agreements in many countries.[22] These rewards depend on the value of the patent to the company, and patent citations are a clear marker for the economic value of patents (Trajtenberg, 1990; Hall et al., 2001).

Second, and most relevant for our purposes, there can be indirect compensation for an inventor's ability. An employer could pay a higher salary or promote star innovators, with a stellar track record of patent quality, in order to both attract and retain crucial talent with the best patenting and innovation ability (Chesbrough, 2006).[23] For a lot of companies, patent licensing is also a major source of revenues, justifying the need to hire the best innovators. For instance, IBM collects more than $1 billion in licensing revenues (Ryder and Madhavan, 2014). The importance of "stars," who play a crucial role in the formation or transformation of many industries, has been emphasized by Zucker and Darby (2014). Note that we are not just trying to measure the income flow from any given patent but rather to proxy for an inventor's full earnings using quality measures based on his patents.

Link between income and citations: Whether it arises from the direct or indirect channel, there appears to be a strong link between the value and quality of patents and the inventor's income. The distribution of rewards for patents seems to be highly skewed towards high quality inventors. Toivanen and Väänänen (2012) find that Finnish inventors receive a temporary reward equal to 3% of earnings for any patent grant. This hides important heterogeneities based on patent quality: moderately cited patents (with 20 to 30 citations) generate a premium of around 20% in annual earnings, while highly cited patents receive an earnings premium of 30% three years after the grant. Harhoff and Hoisl (2007) use data for Germany, where the employer has the right to claim the invention and, if he does, needs to reasonably compensate the employee in proportion to the value of the invention. They also find that the share of the salary received as a compensation for an invention is highly skewed with a few top inventions doubling the inventor's salary. Top 5% inventors receive a 20-50% increase in their salary per invention. Similarly, as a compensation for all inventions, the top inventors' salaries can be multiplied by a factor of 5. Giuri *et al.* (2007) find in the PatVal European inventors survey that 42% of inventors receive a monetary award for their patents, and that for 4% of the respondents, these monetary rewards are permanent. In Swedish administrative data, numbers provided to us by Olof Ejermo show that the most cited inventors earn 60% more than the less cited ones.Bell *et al.* (2015) find using administrative data covering the population of patent applicants in the U.S. that the distribution of income is highly skewed towards superstar inventors with many citations. In their data, the mean income at ages 40-50 is close to $700,000 for inventors with 300 citations, and around $350,000 for inventors with 100 citations.

[22] From our own calculations and reading of the legal rules, 14% of patents in our data come from legislations where by law the employer owns the patent, 30% come from legislations where the employee owns the patents, and the rest come from legislations where ownership is determined by contractual agreements.

[23] Chesbrough (2006) states that "R&D managers often use the number of patents generated (...) as a metric to judge the productivity of (...) [a] person or organization."

As described in detail in Section 5.1, we will allow the ability of the inventor to be rewarded differently in different countries and we will introduce ability and country specific trends in compensation. Additional characteristics used to control for the counterfactual wage are described there as well (such as quality measures, technological field, goodness of fit with the destination, tenure/experience, ability-technological-field-country-specific trends, etc.).

Step 2: Using the effective top marginal tax rate. We use the effective top marginal tax rate as our tax measure.[24] The estimate obtained is not necessarily interpretable as "the migration elasticity to net-of-tax income" for several reasons. First, the average tax rate is not equal to the marginal tax rate because the tax system is not linear. Nevertheless, the top marginal tax rate is likely a good approximation to the average tax rate for top earners.[25] In addition, conditional on being in the top tax bracket, the top tax rate is exogenous to earnings, unlike the average tax rate. The estimate is also interesting *per se* since the top marginal tax rate can also be viewed as a "success tax."

Second, the estimated response will combine firm and worker responses (more on this in Section 5.4) and will be a function of institutional features (e.g., visa regulations, as illustrated starkly for the case of Russian inventors in Section 4.1), both of which prevent an easy mapping from the reduced form estimates to behavioral primitives.

Finally, there are other taxes which may influence migration decision, such as corporate taxes or capital gains taxes. However, we limit the sample to inventors who are employees and who hence receive the bulk of their income as ordinary personal income. These inventors innovate within companies and, typically, their employers are the owners of the patents obtained. To a first order, this allows us to abstract from other forms of taxation such as capital taxation, corporate taxation, or royalties' taxation, and instead focus on personal income taxation. We do control for capital gains and corporate taxes in Appendix Table A14 and check our results on the full sample, including non-employees in Appendix Table A16.

There are certainly some complications with foreign tax rules and regimes across different countries, which we are not able to account for given our data. For instance, an inventor living temporarily in the UK but domiciled abroad can choose to some extent how to be taxed on his income earned abroad (on an "arising basis" or on a "remittance basis"). Depending on the inventor's (unobservable) legal arrangements and future plans, this might lead to somewhat different effective marginal tax rates. In the analysis, we assume that, to a first order, the inventor pays the taxes of the country he physically resides in. The exception is for U.S. inventors who are taxed on

[24] One way to interpret the obtained elasticity is as a reduced form estimate, where the top marginal tax rate can act like an instrument for the top average tax rate. The implicit "first-stage" is indeed significant: from our own computations, we see that changes in the top marginal tax rate are very strongly correlated with changes in the average tax rate on the top 1% of the income distribution.

[25] Reassuringly, Kleven, Landais, and Saez (2013) show that the elasticities obtained for football players using the marginal top tax rate versus actual average tax rates are very similar. This hinges on the fact that those football players considered are well above the top tax bracket in terms of earnings. In our data we find a very strong correlation between the average tax rate on top earners (evaluated using the tax codes for different countries) and the top marginal tax rate.

their worldwide income.[26]

The effective top marginal tax rate is computed including all relevant taxes on labor income: the individual local, state, and national tax rates, the uncapped payroll taxes, and value-added taxes. These series come from Piketty, Saez, and Stantcheva (2014). For U.S. citizens, who are taxed on worldwide income, a special top tax rate is computed for each possible location choice, taking into account the Foreign Tax Credit formula.[27] We drop people in the years in which they are observed in different countries within the same year, as it is not clear what tax rate their yearly income was subject to in those years.[28]

Step 3: Identification using different quality inventors. As Appendix Figure A1 shows, there have been many, both small and large, top tax rate changes in our sample. The simplest identification strategy exploits these variations across countries and time (i.e., country-by-year variation) in top marginal tax rates. In this case, it is still important to control for country fixed effects, year fixed effects, as well as for country specific linear trends to filter out longer-term country-specific evolutions in innovation, incomes, and migration. We view this strategy only as a first pass, as there may be other factors correlated with top marginal tax rates which vary at the same time in a given country, such as the business friendliness of the environment or research stimulating policies. If there are general equilibrium effects of top taxes at the country-year level, these will also be loaded on the estimated coefficient of top taxes.

Our main and preferred identification instead filters out all variation at the country-year level and exploits the differential impact of changes in the top marginal tax rate on the top superstar inventors and slightly lower quality inventors. The idea is that the top 1% inventors and those of slightly lower qualities should be subject to the same country-year level policies and economic effects because they are all very high quality inventors. However, only inventors who are actually in the top bracket are directly affected by the changes in top tax rates. The evidence presented below suggests that top 1% superstar inventors are well in the top tax bracket. While the top 1-5% are still very likely to be in the top bracket, the likelihood of being in the top tax bracket declines sharply as we move down through the top 5-10% to the top 10-25% or the below top 25% of inventors.[29]

Conceptually, this main identification strategy is akin to a "fuzzy regression discontinuity design" where the quality ranking generates a fuzzy threshold above which inventors are "treated" by the top tax rate changes. The lower quality, top 5-10%, top 10-25%, or below the top 25% groups serve as control groups for the top 1%. The top 1-5% is treated as a buffer group, and not as a

[26] Incidentally, in Section 6.3, we consider long-term mobility, which potentially allows a clearer equivalence between geographical location and tax residency, since it is harder to shift income abroad when residing long-term in a different location.

[27] Given the Foreign Tax Credit rules for U.S. citizens, we set the tax rate for U.S. citizen abroad equal to the U.S. tax if and only if the foreign tax rate is smaller than the U.S. tax rate: this was frequently the case before 1985 for the 8 countries under consideration, but not the case anymore after 1985.

[28] Less than 0.2% of the observations are dropped for this reason.

[29] As explained in Section 5.1, the "below top 25%" group refers to inventors who have been or will be in the top 25% at some point during their life in the sample, but are not currently in the top 25%.

good control group, since inventors in this group are too likely to also be treated.

The discontinuity is fuzzy because tax brackets differ across countries and because we do not know counterfactual income and have to instead use quality rankings. We hence establish in two steps, by reference to the literature, that, first, quality is highly positively related to income and that, second, inventor's income distributions are such that the top income inventors are in the top bracket (this is done further below). A useful way to think about this is that the quality groups capture the propensity to be "treated" by the top marginal tax rate.

When choosing the control group, there is a typical trade-off between maximizing comparability (i.e., choosing a group very close in quality e.g., the top 5-10%) and minimizing the control group's propensity of being treated (i.e., choosing a lower quality group such as the below top 25%). For instance, the top 1% (the treatment group) and the top 5-10% are likely extremely comparable groups along all dimensions. However, the top 5-10% may still include quite a few "treated" inventors in the top bracket. In addition, inventors in the top 5-10% may still feel the indirect motivational effect of the "success tax." As a result, we may underestimate the "treatment" effect of the top tax rate if comparing only to the top 5-10% group. We therefore find it useful to show the estimated coefficients for all inventor groups and to compute the elasticities for different choices of control group. We hence include in the regression the full set of interactions of the retention rate with indicator variables for being in the top 1%, the top 1-5%, the top 5-10%, the top 10-25%, and the below top 25%. To reiterate, this identification allows us to include country-year fixed effects in the estimation in order to filter out other contemporaneous changes in the country.

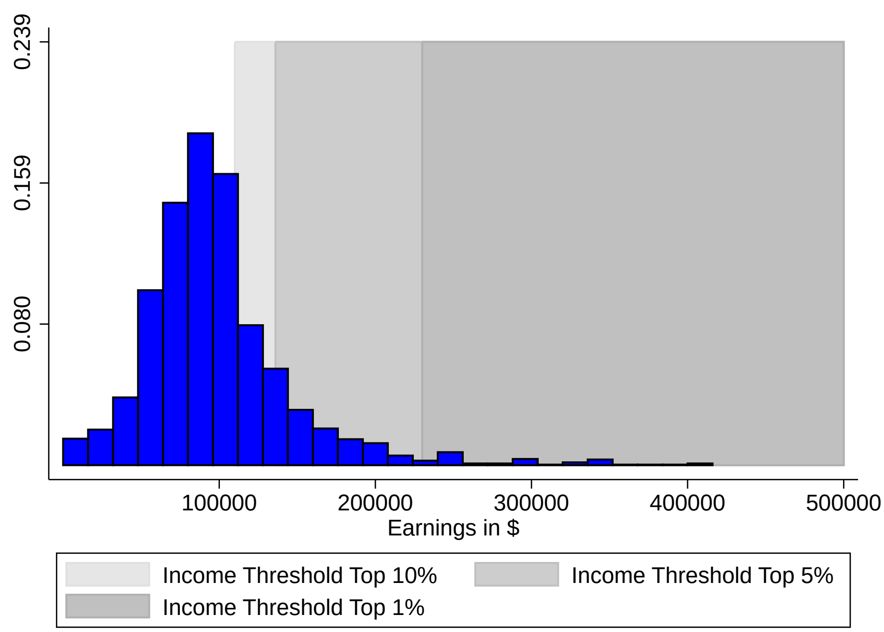
Our strategy is similar in spirit to that in other studies which exploit the differential impact of tax changes across income tax brackets, using lower income brackets as control groups for top brackets (Eissa, 1995), or across households with different non-labor income and family size (Eissa and Hoynes, 2006). We are most closely inspired by a strategy that captures the differential propensities of being "treated," as in Hoynes and Schanzenbach (2009) or Hoynes and Schanzenbach (2012) who multiply Food Stamp program introduction by propensity to participate (proxied by group-level food stamp participation): in our case, we interact the top tax rate with the inventor's quality ranking (which proxies the propensity of being in the top bracket). Note that we are unable to compute an actual propensity to be treated as in Hoynes and Schanzenbach (2009). The problem for doing this is bigger than just the fact that we do not have data on inventor income distributions conditional on quality: we do not know *counterfactual* income (and there is no dataset that contains it) and so we would not be able to know the inventor's actual propensity to be treated in each potential destination country in each year.

For this identification, we may worry that changes in top tax rates are correlated with changes in the average tax rate of the slightly lower quality groups. Historically, this does not seem to have been the case in our setting. There have overall been 218 top tax rate changes in our sample. For only 7 of these have the top marginal tax rate and the tax rate in the bracket below both increased, and for only 14 have they both decreased. Note here that we would mostly be worried about situations in which the top marginal tax rate and the rate in the bracket below would move in the opposite direction, which is the only case that would lead us to overestimate the effect of the top MTR with this identification strategy. Fortunately, this has never occurred in our sample.

Inventors' income distributions: Because of our identification strategy, we need to check that the top quality inventors are in the top income tax bracket, while the propensity of being in the top bracket declines for lower quality inventors. This is done in two steps. First, recall that we established a very significant positive link between citations and income above. We now focus on the income distribution of inventors – keeping this link in mind– to show which percentiles of the inventors' income distribution are in the top bracket and which ones are below.

Using data from the 2003 National Survey of College Graduates public-use microdata from the NSF (NSF, 2003), Figure 2 shows that 44% of inventors are in the top 10% of the U.S. income distribution, 18% are in the top 5%, and 1% are in the top 1% of the U.S. income distribution. A more reliable source is the administrative data used in Bell et al. (2014) which shows that, in terms of income, the top 1% inventor earns \$1.6 million and the top 5% earns \$500,000. The median inventor earns \$114,000 and the mean inventor earns \$192,000. The top 1% highest quality inventors are hence quite likely to be very high up in the income distribution. The top 1-5% inventors by quality are still very high up and have a large propensity of being in the top bracket. The top 5-10% by quality have a lower propensity.

Figure 2: Distribution of Inventor Earnings in the NSF Survey 2003



Notes: The data is from the 2003 National Survey of College Graduates public-use microdata from NSF (NSF, 2003). The earnings represented are those of the college graduates with at least one patent and who report being currently employed. The sample size is 3142.

Turning to European countries and Japan, Figure 3 uses inventor survey data, and plots the income distribution relative to the top bracket. It highlights that the top 1% of inventors in terms of income are clearly comfortably in the top income tax brackets, and the top 1-5% or the top 5-10% are still quite likely to be in the top tax bracket. Below the top 10% the propensity to be treated seems sharply lower.[30]

3 Stylized Macro Facts

We start by providing some stylized macroeconomic facts about the correlations between top retention rates and migration at the country-year level. This evidence is suggestive that, even at an aggregate level, there is a significant correlation between inventor migration and top taxes, which is concentrated on top quality inventors.

3.1 Macro correlations in the inventor data

Figure 4 considers the relation between migration and top tax rates where each dot represents one country in a given year. In all panels, the outcome variable is adjusted for a country's GDP, patent stock, country fixed effects and year fixed effects, thus filtering out time invariant cross-country variation.[31]

Panel A focuses on the fraction of domestic inventors who remain in their home country. Figure 4a shows that the decision to remain in the home country is significantly affected by top tax rates, with an elasticity of 0.08. On the other hand, there is no significant relation between top tax rates and the fraction of low quality inventors in Figure 4b as should be expected given that these inventors are not in the top tax bracket. Panel B turns to the number of foreign inventors as a fraction of all inventors in a country. Figure 4c again shows that top quality foreigners exhibit a significant elasticity of 0.47, while the elasticity of low quality foreigners in Figure 4d is not significant. Table 2 columns (1) and (2) summarize these elasticities.

3.2 Cross-check using the PCT data

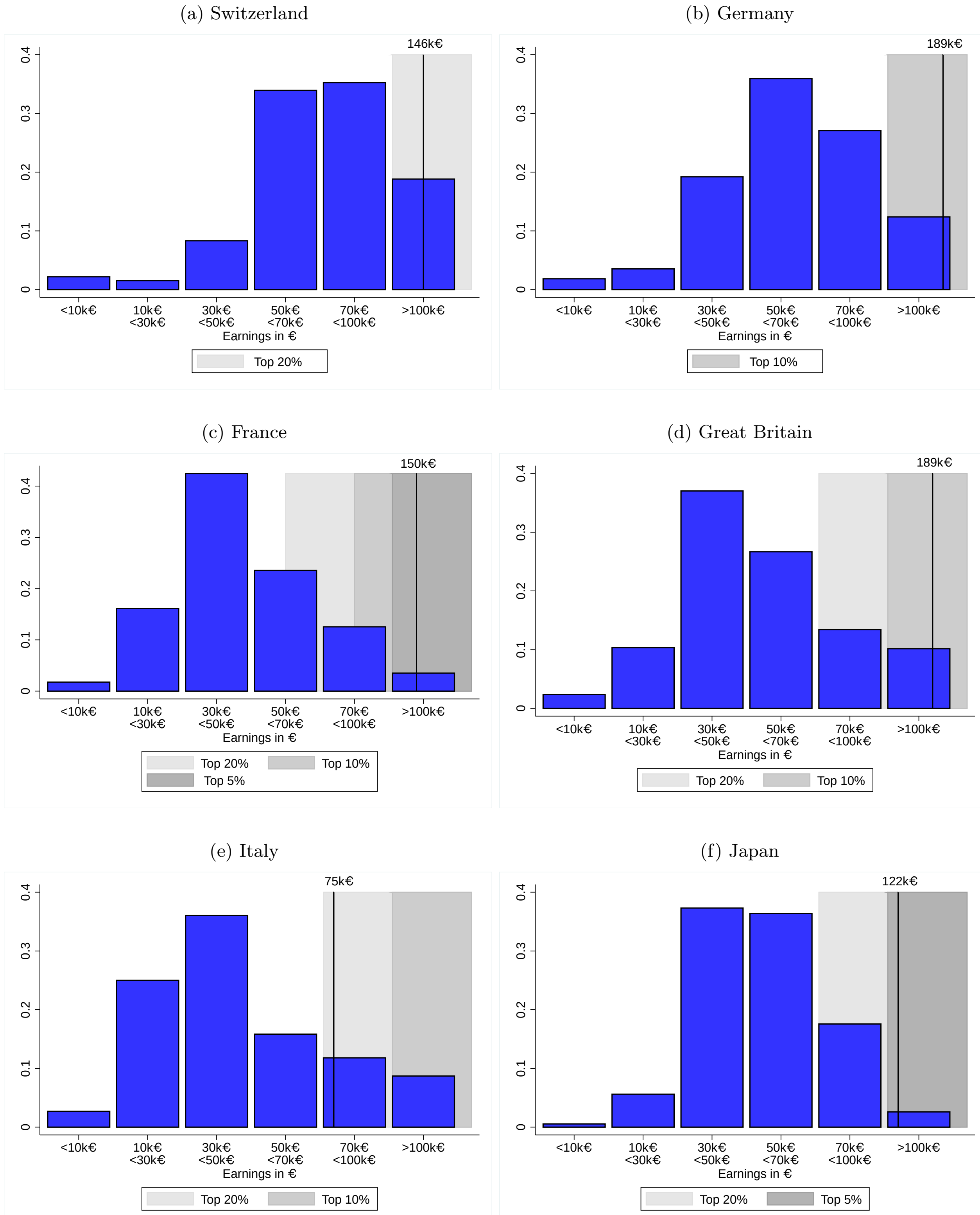
As an additional piece of macro evidence we use the alternative data source from patents filed under the Patent Cooperation Treaty (PCT), described in Section 2.3.

Figure 5a plots the average share of domestic inventors in each country against the average top retention rate over the period 1980-2004. There is a significantly positive relationship, with an

[30]In Swedish administrative data matched to the patent data, even more lower quality inventors seem to be treated. We can see that in 2000, the top 10% inventor earns SEK 648,400, the top 5% inventor earns SEK 775300 and the top 1% earns close to 2 times more at SEK 1,171,000. The median inventor earns SEK 355,200. The top tax bracket threshold in that year was SEK 374,000. Olof Ejermo kindly provided us with these numbers from the Swedish administrative tax data.

[31]In these partial residual plots, for any outcome Y we regress $\log(Y)$ on country GDP, patent stock, country fixed effects, year fixed effects, and the top log retention rate, clustered at the country level. We then construct the adjusted outcome as $\log(Y)$ from which we subtract all covariates (except the log retention rate) times their estimated coefficients.

Figure 3: Income Distributions of Inventors

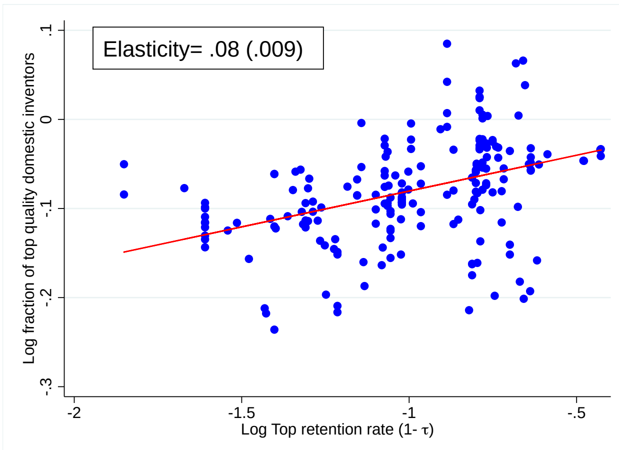


Notes: Survey data from Gambardella *et al.* (2014). The number of respondents for each country are: Switzerland: 457, Germany: 3403, France: 1307, Italy: 966, Great Britain: 551, Japan: 2927. The black vertical line represents the top income tax bracket threshold for individuals. The bins are constrained by the data we have available: these are the income brackets that the survey asked about. We are hence unable to see the exact cutoffs for all the percentiles we are interested in.

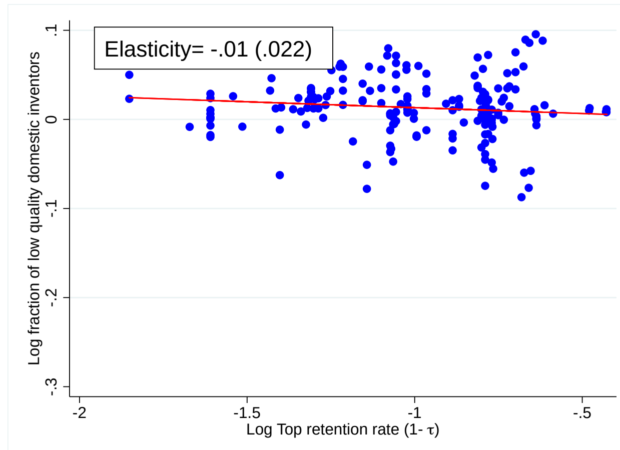
Figure 4: Top taxes and % of domestic and foreign inventors 1977-2000

Panel A: Fraction of domestic inventors in home country

(a) Top quality inventors

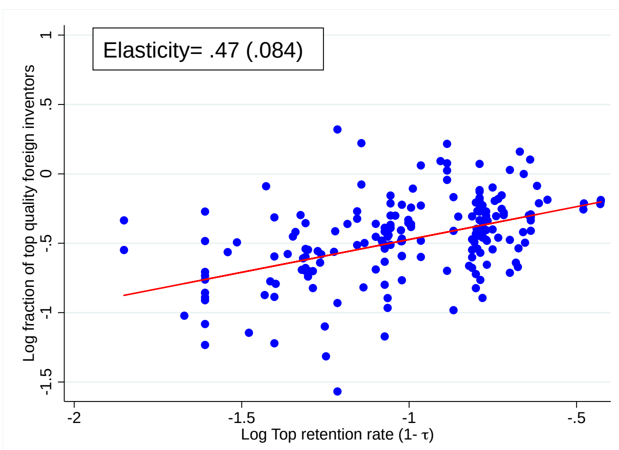


(b) Low quality inventors

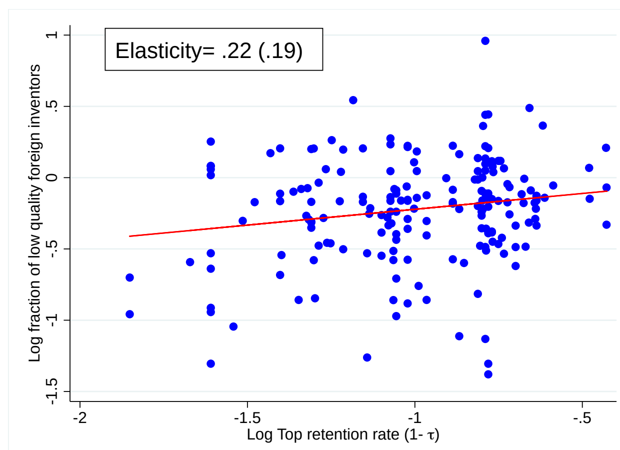


Panel B: Fraction of foreign inventors

(c) Top quality inventors



(d) Low quality inventors



Notes: Each outcome variable at the country-year level is regressed (in logs) on the country's patent stock, GDP per capita, country fixed effects, year fixed effects, and the log retention rate, weighted by the number of inventors in that country and year. The elasticities are reported in each panel with standard errors clustered at the country level. Each dot represents the adjusted log outcome variable, namely the log outcome from which we subtract all covariates (except the log retention rate) times their estimated coefficients. Regression lines are depicted in red. In panel (a), the outcome is the fraction of top 25% inventors working in their home country (number of top 25% inventors working in their home country divided by the total number of top 25% inventors from that country). Panel (b) considers the fraction of low quality bottom 50% inventors who work in their home country. Panel (c) considers the fraction of top 25% foreign inventors (the number of top 25% foreign inventors over the number of all inventors residing in the country). Panel (d) considers the fraction of low quality bottom 50% foreign inventors. Top retention rates are significantly correlated with the shares of top quality domestic and foreign inventors, but not with the shares of lower quality inventors.

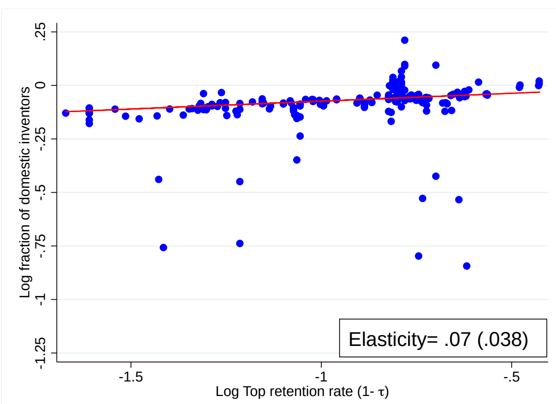
Table 2: Macro elasticities by inventor quality

	Benchmark DID		PCT
	Top quality inventors	Low quality inventors	All inventors
	(1)	(2)	(3)
Domestic Elasticity	0.080***	-0.013	0.074*
	(0.009)	(0.022)	(0.038)
Foreign Elasticity	0.473***	0.222	0.984*
	(0.084)	(0.190)	(0.483)
(Domestic) Observations	192	192	244
(Foreign) Observations	191	188	238

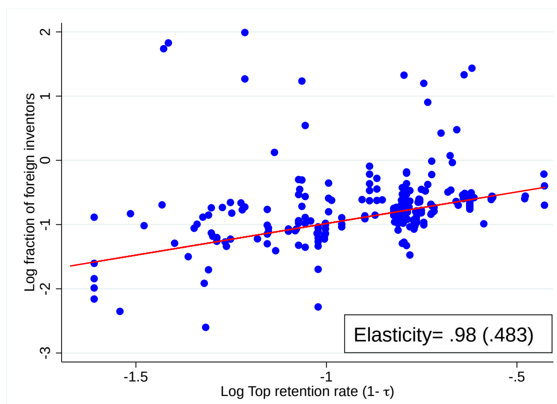
Notes: The table reports the elasticities of the number of domestic and foreign inventors to the top net-of-tax rate. Each outcome variable is at the country-year level. We regress the log outcome on the log top retention rate, the country's patent stock, GDP per capita, country fixed effects, and year fixed effects, weighted by the number of inventors in that country and year. Standard errors clustered at the country level are reported in parenthesis. Columns (1) and (2) correspond to the outcomes from the DID, as represented in Figure 4. Column (3) corresponds to the PCT data from Figure 5. The "domestic elasticity" is the elasticity of the number of domestic inventors who remain to work in their home country. The "foreign elasticity" is the elasticity of the number of foreign inventors who work in a country. As confirmed by the micro level results in Section 5, top quality inventors are sensitive to top tax rates and the foreign elasticity is larger than the domestic one. $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

Figure 5: Top Taxes and % of domestic and foreign inventors in the PCT

(a) Fraction of domestic inventors in home country



(b) Fraction of foreign inventors



Notes: In both panels, the outcome variables at the country-year level are adjusted for the country's patent stock, GDP per capita, country fixed effects, year fixed effects, as described in the note to Figure 4. The elasticities are reported in each panel with standard errors clustered at the country level. The regressions are weighted by the number of inventors for each country-year observation. In panel (a), the outcome is the fraction of inventors working in their home country divided by the total number of inventors from that country. Panel (b) considers the number of foreign inventors over the number of all inventors residing in the country.

elasticity that is extremely close to the one obtained in the DID. Figure 5b confirms that foreign inventors exhibit a much higher elasticity than domestic inventors.

4 Country Case Studies

We now turn to specific country case studies which exploit large historic breaks due to migration or tax reforms.

4.1 Russian inventors' migration and the Soviet Union collapse

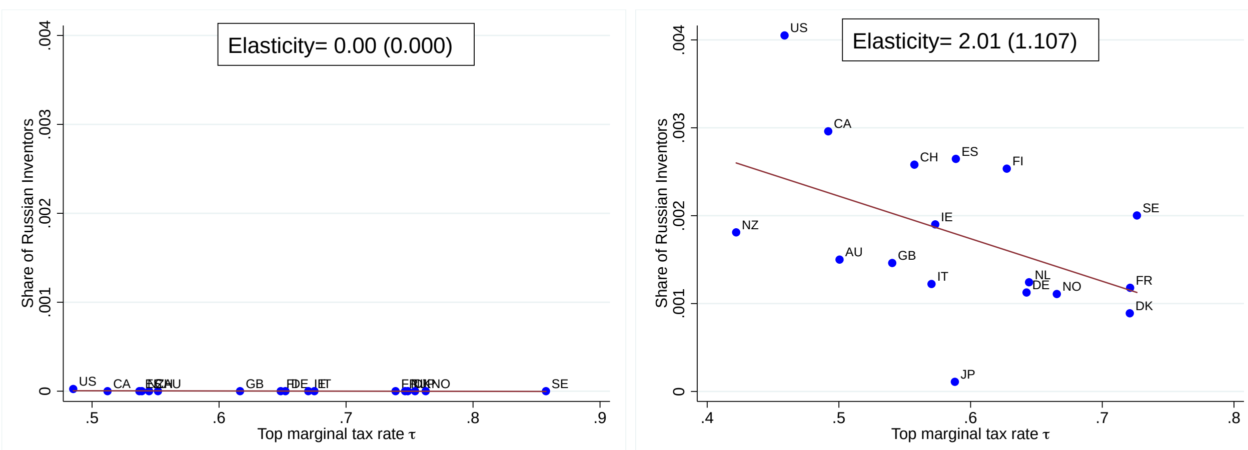
Our first case study concerns Russian inventors, a group of inventors whose migration was essentially impossible before the collapse of the Soviet Union in 1991.

We use two sources of data, starting with the PCT data which has actual nationality information. As a first pass, Figure 6, plots the average number of Russian immigrants in different countries as a function of the average top retention rate before and after the collapse of the Soviet Union. Before 1991 (in panel 6a), the relationship is flat at zero, since Russian nationals were not able to react to tax differentials because of stark migration restrictions. By contrast, after 1991 (in panel 6b), a significantly negative relationship between Russian inventors and top tax rates rates appears.

Figure 6: Russian inventors' migration in the PCT data

(a) Pre Soviet Union collapse

(b) Post Soviet Union collapse



Note: Based on the PCT data that contains nationality information. Each dot represents one country's average outcome over the period. AU= Australia, AT=Austria, BE=Belgium, DK=Denmark, UK=England, FR=France, DE=Germany, FI= Finland, GR=Greece, IE = Ireland, IT=Italy, JP = Japan, NL=Netherlands, NO=Norway, ES=Spain, SE=Sweden, CH=Switzerland. Panel (a) is the pre-Soviet Union collapse period (1977-1991) and panel (b) is the post-Soviet union collapse period (1992-2003). The elasticities reported (with standard errors in parenthesis) come from an OLS regression of the log outcome on $\log(1-\tau)$ where τ is the top marginal tax rate on the x-axis (for the pre-period, the number of Russian inventors is zero so we use the level of the outcome in the regressions). When migration was allowed post 1991, a negative correlation between Russian migration and top tax rates appeared.

Next we move beyond correlations and exploit our main identification strategy, which consists in comparing the effects of the top tax rate on top quality relative to lower quality inventors. For this, we turn to our main DID, which has detailed quality measures for inventors. We combine the

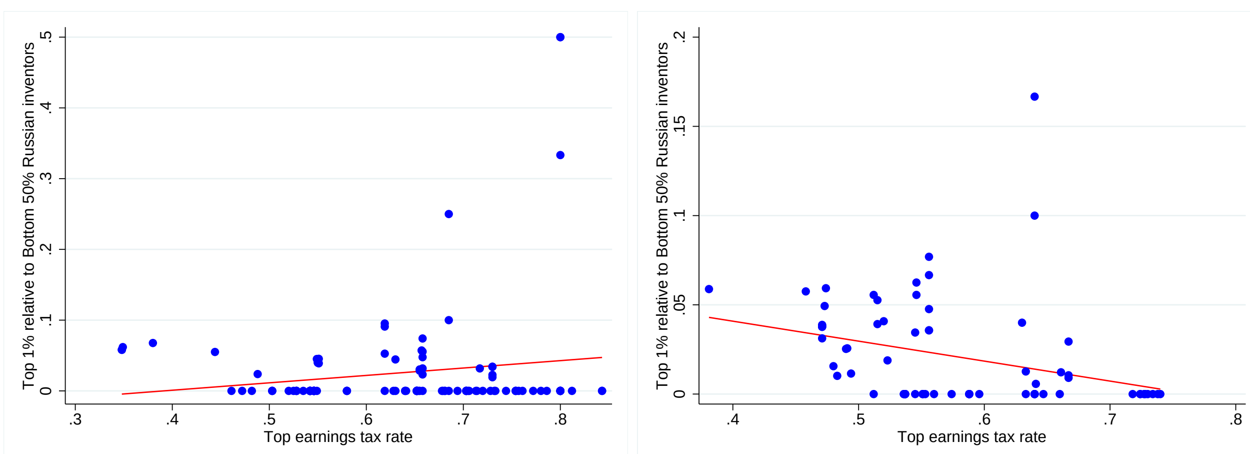
DID with the ethnicity data by Kerr (2008), to circumvent the problem of the lack of nationality information. Kerr (2008) uses an ethnic names database to assign inventors' ethnicities, so we know which inventors have Russian ethnicity. Hence, there are of course inventors of Russian ethnicity abroad even before 1991 (unlike for the pure nationality measure from the PCT).

We can now exploit the differential impact of the top tax rate on higher versus lower quality inventors. Low quality Russian inventors can here serve as a control group that shares similar affinities with each potential destination country as higher quality inventors, and would hence filter out reasons unrelated to taxes (but potentially correlated with them) for which Russian inventors may want to move to specific countries. Low quality Russian inventors, however, should not be directly affected by top tax rates.

In Figures 7a and 7b, we plot the ratio of foreign top 1% inventors over bottom 50% inventors as a function of the top tax rate, pre and post 1991. Each dot represents a country-year observation. We can see the clear differential effect of the top tax rate on top quality Russian inventors relative to low quality Russian inventors. Pre-1991 there is a slightly positive (insignificant) relation between the relative number of foreigners in the top 1% versus the bottom 50% and top tax rates, but this relation becomes negative and significant post-Soviet Union collapse.

Figure 7: Top quality versus low quality Russian inventors' migration

(a) Top 1% over Bottom 50% inventors pre-1991 (b) Top 1% over Bottom 50% inventors post-1991



Note: Based on DID, combined with ethnicity data from Kerr (2008). Kerr (2008) uses an algorithm that identifies Russian ethnicity inventors. The figures represent the ratio of top 1% superstar Russian inventors relative to the bottom 50% low quality Russian inventors. Each dot represents a country-year observation. Before 1991 there is some ethnic Russian diaspora abroad, which is not significantly correlated with top tax rates (coefficient of 0.105, with standard error 0.09). Post Soviet Union collapse, there is a significant negative correlation between superstar top 1% Russian inventors relative to lower quality Russian inventors and top tax rates (the coefficient is -0.111, with standard error 0.028). Analogous regressions with additional controls for log GDP per capita, log number of patents and country and year fixed effects produces coefficients for pre-1991 of -0.09, with standard error 0.075, and for post-1991 of -0.199, with standard error 0.092.

Table 3 shows the elasticity estimates for inventors of different quality. It confirms that while top 1% superstar inventors' migration was unsurprisingly not related to top tax rates before the collapse of the Soviet Union, there is a very significant elasticity to top tax rates post 1991. The same effect, but much weaker, is visible for top 1-50% inventors. The elasticity of the top 1% superstar Russian inventors matches quite closely the elasticity of foreigners in the micro estimation in Section 5.[32]

Table 3: Elasticity of Russian inventors to top retention rates

	(1) Top 1%	(2) Top 1-50%	(3) Bottom 50%
Pre Soviet Union collapse	0.0878	0.0779	0.368**
	(0.193)	(0.131)	(0.143)
Post Soviet Union collapse	1.154***	0.398**	0.347*
	(0.263)	(0.191)	(0.186)
Observations	192	192	192

Based on DID, combined with ethnicity data from Kerr (2008). OLS regression of the log number of Russian inventors in a country each year on the log of the top retention rate. Each column reports results for a subset of inventors, namely those in the top 1%, the top 1-50%, and bottom 50% as ranked by the benchmark measure of citations-weighted patents computed according to formula (1). The coefficients are the elasticities to the top retention rate of Russian inventors in each quality group. All regressions control for year fixed effects, country fixed effects, GDP per capita, and country's patent stock. Robust standard errors are in parentheses. $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$

4.2 United States Tax Reform Act 1986

Next, to dig deeper into the identification of the effects of top tax rates, we exploit the quasi-experimental variation provided by two large tax reforms: the U.S. Tax Reform Act of 1986 and Denmark's 1992 preferential tax scheme for foreigners.

To study the effects of these reforms on inventor migration, we use the synthetic control method by Abadie et al. (2010). It consists in building a synthetic control country that is a weighted average of other countries in the sample. The weights are chosen to minimize the pre-reform distance between the country under consideration (i.e., U.S. or Denmark) and the synthetic country along dimensions of interest. The synthetic country plausibly represents what would have happened in the control country absent the reform. Appendix D provides the details.

To estimate the elasticities to top tax rates, we regress at the country-year level the log of each outcome variable on the log retention rate, a dummy for the post-reform period and a dummy for country (there are two "countries" here, namely the treated country and the synthetic control country), where the log retention rate is instrumented by the interaction of the country dummy and the post-reform period dummy.

The U.S. Tax Reform Act of 1986 reduced top marginal tax rates from 50% to 28%.[33] The corresponding change in average tax rates was very concentrated at the top of the income distri-

[32]Note that because Kerr (2008) assigns ethnicity (not nationality), it is expected to see some ethnic Russian inventors abroad in the Soviet Union era. What matters here is how the elasticity changed in the post collapse period. For low bottom 50% Russian inventors, as expected, the correlation between migration and top tax rates did not change post-Soviet Union collapse.

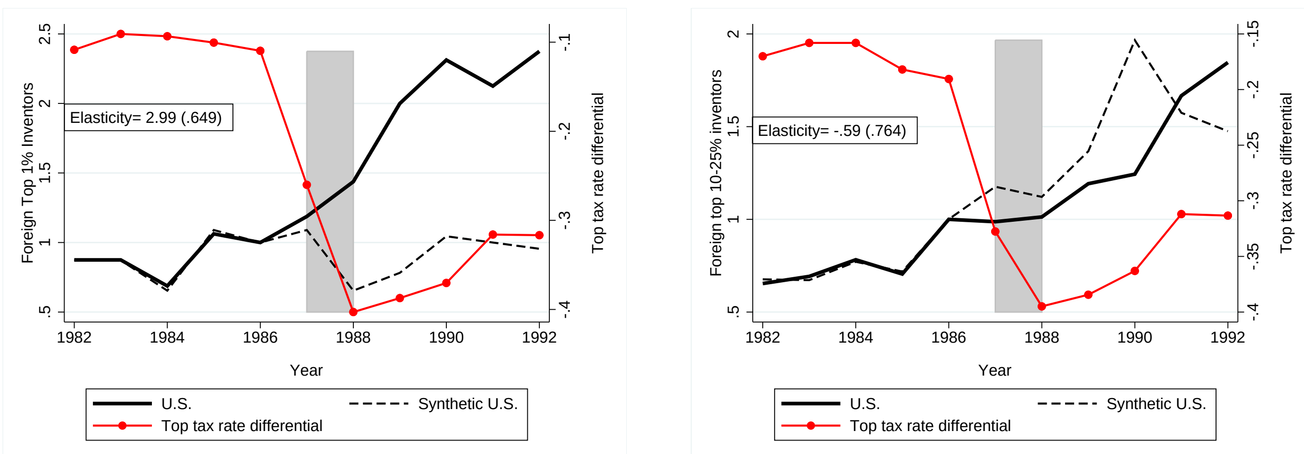
[33]There is still a debate about the extent to which the decrease in marginal tax rates was offset by base broadening, which was one of the goals of the reform. Gravelle and Hungerford (2012) argue that the tax base of the individual income tax was not widely broadened. Bakija and Steuerle (1991) find an overall modest increase in the tax base of 3.8%.

bution. According to our own calculations, average tax rates decreased by a large 7.6% for the top 1%, decreased by only 2% for the top 5% of inventors, and actually increased by 0.7% for the top 10% of the income distribution.

Figure 8 plots the number of foreign inventors in the U.S. and in its synthetic counterpart, normalized by a base year (1986).[34] We again use our preferred identification (and the differential impact of top tax rates on inventors of different qualities), and show the effects of this reform separately for top quality (top 1%) and slightly lower quality (here, top 10-25%) inventors.[35]

Figure 8: The 1986 Tax Reform Act and Foreign Inventors in the U.S.

(a) Foreign Top 1% inventors (b) Foreign Top 10-25% inventors



Notes: The 1986 Tax Reform Act (implemented during the fiscal year represented by the grey vertical area), lowered top tax rates in the U.S. In each panel, the fraction of foreign inventors in the U.S. versus the fraction of foreign inventors in the synthetic control country are depicted. The weights on countries forming the synthetic country are set to minimize the distance to the U.S. in the pre-reform years (see the text and Appendix D for details). Panel 8a shows the number of top 1% foreign inventors in the U.S., while panel 8b displays the number of top 10-25% foreigners in the U.S. Both are normalized by a base year (1986). The red line shows the top tax rate differential between the U.S. and the synthetic control country (on the right axis), i.e., $\tau_{US}/\tau_{\text{synthetic}} - 1$. We report the difference-in-difference elasticity estimates described in the text.

Indeed, the differential effect of the top tax rate change on the foreign superstar top 1% inventors is visible in two ways. First, purely exploiting the logic of the synthetic control method, we can show that the number of foreign top 1% superstar inventors increased drastically relative to a scenario with no reform (as proxied by the synthetic control), while this was not the case for the lower quality top 10-25% foreign inventors. In panel (a), the red line (on the y axis) is the top tax differential between the U.S. and the synthetic country, defined as $\frac{\tau_{US}}{\tau_{\text{synthetic}}} - 1$. While top tax rates

[34]Some normalization is necessary because the U.S. is the largest country in the data and it would not make sense to try and match levels. It would also not make sense to look at the fraction of foreigners, since the reform was not preferentially targeting foreigners (unlike Denmark's reform discussed in the next Subsection).

[35]Recall that the top 1-5% group still has a very high propensity of being treated and is hence not a good control group. The results described next naturally also hold if we consider even lower quality inventors, e.g., the below top 25%, etc., subject to the trade-offs involved in choosing the control group, as described in Section 2.4.

were about the same in the U.S. and in the control country, they diverge sharply post reform.[36] The pre-reform trends are extremely close in the U.S. and the synthetic country. However, the series for top 1% foreign superstar inventors diverge exactly at the time of the U.S. tax reform between the U.S. and the synthetic control country. When we repeat the analysis for the lower quality group of inventors (the top 10-25%), we do not see such an effect: there is no break in the trend post reform and, if anything, the U.S. series for the top 10-25% foreign inventors fall a bit relative to the synthetic control. Hence, the lower quality top 10-25% foreign inventors, fare similarly in the U.S. (which decreased its top tax rates) and in the other countries (which did not decrease top tax rates).

The second way to see the differential effect of the reform on foreign superstar top 1% inventors is to directly consider the structural break in their series relative to that of foreign top 10-25% inventors. Naturally, all the series for foreigners are growing due to the general growth in patents, inventors, and innovation over time, but we can directly compare what happened to the growth rates of the top 1% foreign inventors and of the top 10-25% foreign inventors after the 1986 reform. Comparing the two solid black lines in panels (a) and (b), we see that while the top 10-25% group was growing faster before the reform, it is the top 1% that started growing much more rapidly post reform. More specifically, pre-reform, the top 10-25% series was growing on average at 12.7% per year, while the top 1% was growing at only 6.8% per year during the same period. On the other hand, after the reform, the top 10-25% was still growing at about the same rate (11.3% per year), while the top 1% drastically accelerated to 16.4% per year. As a result, the composition of foreign inventors in the U.S. dramatically changed after the 1986 reform, and the ratio of the top 1% foreigners to the top 10-25% foreigners drastically increased.

4.3 Denmark's 1992 reform

Next, we consider the Danish 1992 reform, studied in Kleven et al. (2013) and Kleven et al. (2014).[37] The Danish tax reform created a preferential tax scheme for foreign researchers and high-income foreigners. Instead of the usual top tax rate (of 60%), foreigners were taxed for 3 fiscal years at a flat rate of 30% for their 1991-1995 income and then at another reduced rate of 25% after 1995. Since our inventors would typically qualify as "researchers," we can study whether they were affected by this tax scheme.

We revert to using the PCT data, since Denmark is essentially not represented in the USPTO data. Recall that in the PCT data, we do not have inventor quality measures since it is not a panel data. Looking at the effect on all inventors is likely to give us a lower bound of the effects of

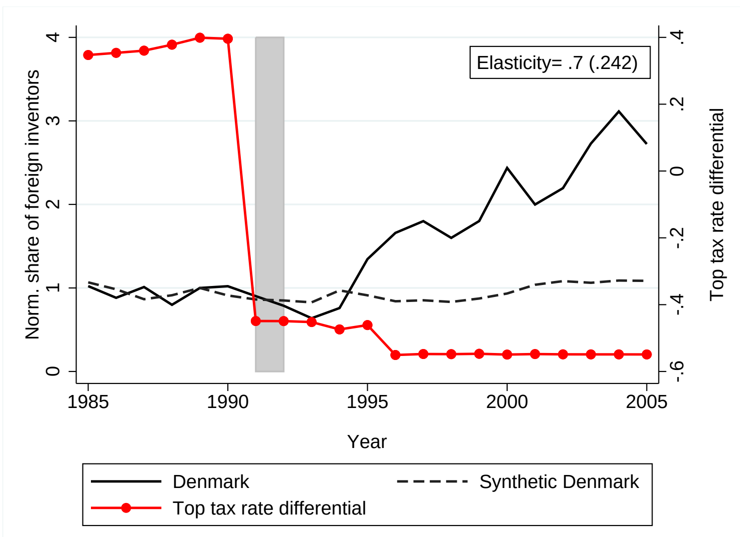
[36] The approach puts significant weight on Canada and on Switzerland for the top quality series and on Canada, Great Britain, and Japan for the lower quality series (see Appendix D and Appendix table A12)

[37] Unfortunately, other such reforms, such as the Spanish law of 2004 studied by Kleven et al. (2013) took place too late given our sample years (and the problem of truncation of patent citations) and cannot be studied well in our setting.

the reform on the high-quality, high-income inventors. In addition, we cannot track who is eligible under the 3-year rule. Overall, this should give us a diluted, or lower-bound effect.[38]

In Figure 9 we plot the fraction of foreign inventors in Denmark and in the synthetic country (again normalized by a base year (1985)), as well as the top tax rate differential on foreigners. We here use as an outcome variable the fraction of foreigners (foreign inventors over all inventors in Denmark), since this allows us to see the differential effect on foreigners (affected by the reform) relative to domestic inventors (not affected by this reform). The synthetic control algorithm puts a large weight on Switzerland and some small weights on Canada and Portugal (see Appendix D), which are the countries in the PCT sample that best match Denmark's pre-reform share of foreign inventors. The effects of the Danish reform on the inflow of foreigners is clear relative to the synthetic control. The implied elasticity is 0.7, which is a bit lower than our preferred elasticity of foreigners in the micro analysis based on the DID in Section 5, but almost identical to the elasticity of foreigners based on the EPO data in Section 7.

Figure 9: Denmark's 1992 Tax Reform and Foreign Inventors



Notes: The 1992 Danish reform (implemented during the fiscal year represented by the grey vertical area), lowered top tax rates for high income foreigners and foreign researchers coming into Denmark for the first 3 years. The figure depicts the fraction of foreign inventors in Denmark versus the fraction of foreign inventors in the synthetic control country (normalized by a base year (1985)). We here use as an outcome variable the fraction of foreigners (foreign inventors over all inventors in Denmark), since this allows us to see the differential effect on foreigners (affected by the reform) relative to domestic inventors (not affected by this reform). The weights on countries forming the synthetic country are set to minimize the distance to Denmark in the pre-reform years (see the text and Appendix D for details). The red line shows the top tax rate differential between Denmark and the synthetic control country (on the right axis), i.e., $\tau_{DK}/\tau_{\text{synthetic}} - 1$. We report the difference-in-difference elasticity estimate described in the text.

[38] The fact that we are obtaining a lower bound effect is confirmed through the comparison to Kleven et al. (2014), who are able to focus on top quality, high earning agents only and find an elasticity of foreigners between 1.5 and 2.

5 Location Choice Model Estimation

In this section, we present the specification of and the results from the multinomial location model estimation that exploits our two identification strategies.

5.1 Specification and Estimation

Utility specification and controls: To model the location choice of inventors, we adopt a multinomial discrete-choice model. Recall from section 2.2 that inventor i in country c at time t obtains utility:

$$U^i_{ct} = u\left(\left(1 - \tau_{ch^it}\right) w^i_{ct}\right) + \mu^i_{ct} \tag{5}$$

For the empirical specification, we assume log utility of consumption (which allows us to rewrite utility as additively separable in pre-tax earnings and the tax rate). Our benchmark sample for the multinomial analysis contains inventors who, over their lifetime in the sample, have ever been or will ever be classified in the top 25% inventors (according to the ranking defined previously).[39] Recall that we restrict the sample to the 8 countries which account for around 90% of all patents in the USPTO. Hence, in each year, inventors face a choice between the United States, Canada, Germany, Great Britain, France, Japan, Italy and Switzerland.[40]

An important part of the analysis is to model as detailed and flexibly as possible i) the idiosyncratic preference component μ^i_{ct} and ii) the counterfactual wage w^i_{ct}. We consider them in turn.

Controlling for the idiosyncratic preference component: The innovator's idiosyncratic preference for any given country at time t is assumed to depend on:

i) Individual-level characteristics $\mathbf{x}_{ti}$: These include the inventor's age, his dynamic quality measure, his technological field, and whether he works in a multinational. In addition, we include indicator variables for whether the inventor is in the top 1%, top 1-5%, top 5-10%, top 10-25%, or below top 25% of inventors as ranked by quality and explained in section 2.3. The benchmark measure used is the citations-weighted patents to date, but we consider all of the measures $q1 - q4$ from formulas (1)-(4). The effect of these individual characteristics is allowed to vary by country (i.e., they are all interacted with country fixed effects). We assign one of six technological fields to the inventor based on the field in which he has most patents. These six fields that come from Hall et al. (2001) include chemical, computer and communications, drugs and medical, electrical and electronics, mechanical and others. Using finer classifications did not change our results.

[39] This is to reduce the computational burden of the multinomial logit with many fixed effects on a very large sample. Each inventor-year observation will be duplicated 8 times for each of the 8 potential country choices, thus drastically scaling up the data size.

[40] Note that the sample cut is based on a static measure of inventor quality, namely those who have *ever* been in the top 25% in their lifetime, while the ranking within that sample is dynamic and can change for a given inventor from year to year. Table 9 considers how the results change if we also rank inventors by the static measures.

ii) Country-level covariates, denoted by $\mathbf{x}_{ct}$: These depend on the identification used. In the case in which we identify off country-by-year variations in the tax rate (identification 1), these covariates include the country's patent stock and GDP per capita, as well as country fixed effects, year fixed effects, and country-specific time trends.[41] For the identification using inventors of different qualities as treatment and control groups (identification 2), we replace the country's patent stock and GDP per capita by country $\times$ year fixed effects.

iii) Controls for inventor-country pairs, denoted by $\mathbf{x}_{cti}$ that capture the goodness of fit of a "match" between country c and inventor i at time t. First, we introduce a home dummy h_{ci} equal to 1 if country c is the home country of inventor i in order to capture a potential home bias. Second, these include the patent stock of country c at time t in the inventor's technological field, which proxies for the quality of professional fit between the inventor and the potential destination country. Third, we include the distance between the inventor's home country and the destination country, and a dummy for whether the home and destination country have a common language. These could all affect the ease of moving.

In Section 5.4 that focuses on the role of companies in the migration decision, we also consider other country-inventor specific covariates such as the share of the innovative activity of the inventor's company in each destination country.

Controlling for the tax rate: As explained in Section 2.3, we use the top marginal tax rate as our tax measure. We include the log of the top retention rate interacted with indicator variables for the inventor's quality rank at time t, denoted by r_{it}, where r_{it} can be the top 1%, top 1-5%, top 5-10%, top 10-25% or below the top 25%. In the case of identification 1, which uses country-by-year variation in the top marginal tax rate, this simply allows for differential effects of the top tax rate on different quality inventors. In the case of identification 2, this captures the propensities of different quality groups of being "treated" by the top tax rate.

We hence obtain the econometric specification for a discrete choice model at the individual inventor and year level. For each potential country choice c:

$$U_{cti} = \alpha_{r_{it}} \log\left(1 - \text{top MTR}_{ct}^{i}\right) + \alpha \log\left(w_{cti}\right) + \beta_c \mathbf{x}_{ti} + \zeta \mathbf{x}_{ct} + \eta \mathbf{x}_{cti} + v_{cti} \tag{6}$$

where $\mathbf{x}_{ct}$ includes GDP per capita, patent stock, country fixed effects and country-specific time trends in the case of identification 1 and country fixed effects and country $\times$ year fixed effects in the case of identification 2. "top MTR_{ct}^{i}" is the effective top marginal tax rate at time t in country c which depends on inventor i: if inventor i is a U.S. citizen, then we apply the foreign tax credit rule for U.S. citizens. Otherwise, it is just the marginal top tax rate of country c at time t without any modification. $\alpha_{r_{it}}$ is a coefficient on the top marginal tax rate that depends on the inventor's quality rank r_{it} (equivalent to interacting the top retention rate with rank dummies).

[41] Note that including year fixed effects does not change anything since the multinomial logit already filters out all variables which do not vary by country.

For our preferred identification 2, to compute the effect of the top tax rate on the top 1% inventors, we need to pick a control group. As explained previously, given the fuzzy design, it is not a priori obvious which group is the best control group. This is why we provide the results for three choices of control groups, the top 5-10%, the top 10-25%, and the below top 25%. Bear in mind that in this estimation sample, the "below top 25%" refers to a still relatively high quality group of inventors, which have been or will at some point be in the top 25%, but are not currently there. If we consider group g as the control group, where $g \in \{\text{top 5-10\%, top 10-25\%, below top 25\%}\}$, then the corresponding effect on the top 1% will be computed as $\Delta\alpha_g = \alpha_{\text{top 1\%}} - \alpha_g$.

Controlling for the counterfactual wage: As explained earlier, a crucial challenge is to control for the counterfactual wage w_{cti} that inventor i would receive in any country c at time t. Part of the wage variation is well absorbed by the aforementioned controls (i)-(iii), which capture aggregate effects at the country-year level and the technological class level. In addition, we consider three more benchmark specifications that progressively add the following, more detailed proxies for $\log(w_{cti})$. Overall, we estimate the following four specifications:

1) Including only controls (i)-(iii).

2) Adding the quality of the inventor at time t, interacted with country fixed effects. This controls for aggregate country-specific effects that vary by ability. Our benchmark measure is, as explained in Section 2.3, citation-weighted patents to date (measure $q1$). Alternative quality measures are considered in the robustness checks in Table 9.[42]

3) Adding country and ability specific trends, i.e., controlling for year trend times country fixed effect times the ability measure, in order to capture differential evolutions over time in the wage premium in different countries.

4) Adding country, ability and technological field specific trends by controlling for year trend times country fixed effects times ability times technological field dummies.

Estimation: Denote by $P^i_{ct} \equiv Prob(U^i_{ct} > U^i_{c't}, \forall c')$ the probability of inventor i to locate in country c at time t. If the error term v^i_{ct} has a type I extreme value distribution, this model can be estimated as a multinomial logit.

Computing Elasticities: To go from the coefficients on the retention rate to elasticities, we follow the computations in Kleven, Landais, and Saez (2013) that we present only very briefly here. In the multinomial model, the elasticity of the probability of inventor i of locating in country c at time t to the net of tax rate $(1-\tau^i_{ct})$, denoted by ε^i_{ct}, is:

$$\varepsilon^i_{ct} \equiv \frac{d\log P^i_{ct}}{d\log(1-\tau^i_{ct})} = \Delta\alpha_g(1-P^i_{ct})$$

where $\Delta\alpha_g$ is the effect, defined above, relative to control group $g \in \{\text{top 5-10\%, top 10-25\%, below top 25\%}\}$. The authors then define the elasticity of domestic players in country c, ε^c_d, and the elasticity of

[42] Note again that the ability measure for an individual is dynamic and changes over life.

foreign players in country c, ε_f^c. Letting I_c and I_c^f be the set of all, respectively, domestic and non-domestic inventors from country c:

$$\varepsilon_d^c \equiv \frac{d\log(\sum_{i\in I_c} P_{ct}^i)}{d\log(1-\tau_{ct})} = \frac{\Delta\alpha_g \sum_{i\in I_c} P_{ct}^i(1-P_{ct}^i)}{\sum_{i\in I_c} P_{ct}^i} \tag{7}$$

Similarly:

$$\varepsilon_f^c \equiv \frac{d\log(\sum_{i\in I_c^f} P_{ct}^i)}{d\log(1-\tau_{ct})} = \frac{\Delta\alpha_g \sum_{i\in I_c^f} P_{ct}^i(1-P_{ct}^i)}{\sum_{i\in I_c^f} P_{ct}^i} \tag{8}$$

Average domestic and foreign elasticities, ε_d and ε_f, then are defined as the weighted average elasticities across all countries:

$$\varepsilon_d \equiv \frac{d\log(\sum_c \sum_{i\in I_c} P_{ct}^i)}{d\log(1-\tau_{ct})} = \frac{\Delta\alpha_g \sum_c \sum_{i\in I_c} P_{ct}^i(1-P_{ct}^i)}{\sum_c \sum_{i\in I_c} P_{ct}^i} \tag{9}$$

Similarly:

$$\varepsilon_f \equiv \frac{d\log(\sum_c \sum_{i\in I_c^f} P_{ct}^i)}{d\log(1-\tau_{ct})} = \frac{\Delta\alpha_g \sum_c \sum_{i\in I_c^f} P_{ct}^i(1-P_{ct}^i)}{\sum_c \sum_{i\in I_c^f} P_{ct}^i} \tag{10}$$

5.2 Results using country-by-year variation

The first identification exploits country-by-year variation in the top marginal tax rate. In this case, the regressions includes country fixed effects, year fixed effects, and country-specific year trends, in addition to the main controls listed above, but no country-year fixed effects. To allow for heterogeneous effects of the top tax rate for different qualities of inventors, the top marginal retention rate is interacted with dummies for the quality ranking.

Table 4 column 1 contains the controls i) - iii) described in Section 5.1. Below the regression estimates, the table reports the values and the standard errors of the elasticities to the net of tax rate of, respectively, the number of top 1% superstar domestic and foreign inventors, as computed from formulas (9) and (10).[43]

The subsequent columns add the remaining controls for the counterfactual wage listed as 2) -4) in Section 5.1: Column 2 introduces the ability of the inventor interacted with country fixed effects. Column 3 adds country and ability specific trends and column 4 adds country, ability and tech field specific trends. The effects of the top tax rate are very stable across the different specifications, which provides some reassurance about the controls for the counterfactual wage.

The top 1% of inventors and the top 1-5% of inventors are significantly sensitive to the top tax rate. The elasticity of domestic superstar top 1% inventors is (in the most detailed specification) 0.024 and that of foreign top 1% superstar inventors is 0.8. Note that these match very well the

[43]The gap between the domestic and the foreign elasticity comes from the multinomial specification, combined with the fact that most inventors remain in their home country (see how the average probability of remaining in the home country scales the elasticities in formulas (9) and (10)). It does however fit very well with the correlations in Section 3 which showed a much higher elasticity for foreigners.

TABLE 4: EXPLOITING COUNTRY-BY-YEAR VARIATION AND GENERAL EQUILIBRIUM EFFECTS

	(1)	(2)	(3)	(4)
Log Retention Rate × Top 1	0.894***	0.895***	0.969***	0.955***
	(0.206)	(0.208)	(0.210)	(0.212)
Log Retention Rate × Top 1-5	0.443***	0.452***	0.522***	0.502***
	(0.133)	(0.134)	(0.133)	(0.134)
Log Retention Rate × Top 5-10	0.148	0.162	0.233**	0.209*
	(0.114)	(0.114)	(0.111)	(0.112)
Log Retention Rate × Top 10-25	-0.123	-0.0987	-0.0222	-0.0460
	(0.0934)	(0.0933)	(0.0890)	(0.0895)
Log Retention Rate × Below Top 25	-0.405***	-0.349***	-0.267**	-0.278**
	(0.113)	(0.119)	(0.119)	(0.120)
Quality× Country FE	NO	YES	YES	YES
Quality × Country FE × Year	NO	NO	YES	YES
Quality×Country FE×Year×Field FE	NO	NO	NO	YES
Domestic elasticity	.02	.02	.024	.024
s.e	(.0047)	(.0047)	(.005)	(.005)
Foreign elasticity	.75	.754	.811	.811
s.e	(.174)	(.175)	(.177)	(.177)
Observations	8644280	8616336	8616336	8616336

Notes: Multinomial logit regressions. Robust standard errors clustered at the inventor level in parentheses. Regressions are based on the disambiguated inventor data (DID) described in Section 2.3 for the period 1977-2000. The data includes inventors located 8 countries: Canada, France, Germany, Great Britain, Italy, Japan, Switzerland, and the United States. All columns contain the following controls also listed in the text. In terms of country-level controls, we include country fixed effects, year fixed effects, and country-specific time trends, country patent stock, country GDP per capita. In terms of country-inventor pair controls, we include a home country dummy, the patent stock of the country in the inventor's technological field, the distance between the inventor's home country and the country, and a dummy for whether the country shares a common language with the inventor's home country. The following inventor-level variables are all included and interacted with country fixed effects: inventor age, technological field of the inventor, a dummy for whether the individual works in a multinational firm. All columns also contain indicator variables for whether the inventor is in the top 1%, top 1-5%, top 5-10%, top 10-25% or below the top 25% of inventors as ranked by quality and explained in section 2.3. Column 2 contains in addition the inventor's citations-weighted patent stock to date (measure $q1$) interacted with country fixed effects. Column 3 contains quality and country-specific time trends. Column 4 contains quality, country and technological field specific time trends. The first row reports the coefficient on the log retention rate interacted with an indicator variable for being in the top 1% of inventors, while the second row reports the standard error. The subsequent row pairs report, respectively, the coefficients from the retention rate interacted with being in, respectively, the top 1-5%, top 5-10%, top 10-25% and below the top 25% of inventors, as well as their standard errors. "Domestic elasticity" is the elasticity of domestic top 1% superstar inventors with respect to the top net-of-tax rate (one minus the top tax rate), while "Foreign elasticity" is the elasticity of foreign top 1% superstar inventors with respect to the top net-of-tax rate. They are computed according to formulas (9) and (10) with $\Delta\alpha_g = \alpha_{\text{top }1\%}$. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

estimates using the below top 25% as a control group in Section 5.3 and Table 5.

It is interesting that the top 10-25% and below the top 25% are actually negatively driven by top tax rate decreases, which can be due to the presence of sorting and general equilibrium effects. Indeed, the lack of country × year fixed effects here means that sorting and general equilibrium effects could be loaded on the estimated coefficients on the log retention rate. The migration of some inventors can have general equilibrium effects and spillovers on other inventors, depending on the structure of the labor market. For instance, if demand for inventors by companies in different countries is rigid, a lower top tax rate could lead to sorting by higher quality inventors, to the detriment of lower quality inventors.[44] A high influx of top 1% and top 5% inventors may displace lower quality inventors.[45]

5.3 Results exploiting the differential effects on inventors of different qualities

Table 5 contains the estimation results for our main identification that compares the effect of the top tax rate on the top 1% inventors relative to lower quality inventors. The coefficient on the retention rate for the top 1% inventors is significantly positive and large. The top 1-5% inventors still exhibit a very significantly positive effect of top tax rates, as can be expected since inventors in this group have a very high propensity of being treated. The top 5-10% inventors also show a positive effect, albeit insignificant, since their propensity to be treated is lower. The coefficients decline monotonically as we move down the quality distribution due to the declining propensities of being treated by the top tax rate.

The lower panel in the table reports the elasticities of domestic top 1% superstar inventors and foreign top 1% superstar inventors for different choices of the control group g, computed according to formulas (9) and (10) for $g \in$ {top 5-10%, top 10-25%, below top 25%}. Depending on the control group chosen, the estimated elasticities of domestic and foreign inventors will be different. As already discussed above, a control group such as the top 5-10% may give a lower bound estimate, since inventors in this group may still be partially treated. A control group such as the below top 25%, while still a relatively high quality group, has a lower propensity of being treated and will hence yield a larger estimated effect of the top tax rate on superstar top 1% inventors. For the control group consisting of, respectively, the top 5-10%, the top 10-25%, and the below top 25%, the estimated domestic elasticities are, respectively, 0.02, 0.02, and 0.03 (and are all statistically significant at the 1% level). The corresponding elasticities on foreign inventors are, respectively, 0.63, 0.84, and 1.04.

For comparison, Kleven, Landais, and Saez (2013) find that, depending on the specification, the domestic elasticity of football players is between 0.07 and 0.16, while the foreign elasticity is between 0.6 and 1.3. The authors suggested that their elasticities might be upper bounds, as football players are highly mobile individuals. In addition, they only consider the European football market, while we also include mobility between different continents.

To roughly illustrate these elasticities, suppose that the average country decreases its top tax

[44] Kleven, Landais, and Saez (2013) also explore sorting effects in a rigid demand model for football players.

[45] Another possible interpretation of the negative coefficient on lower quality inventors, is that there are direct aggregate effects of the top tax rate through tax revenue, which benefits lower quality inventors. However, controlling for total tax revenue per capita in country c and time t leaves the effects of the top tax rate on different quality inventors entirely unchanged.

Table 5: Effect of the top retention rate on innovators' mobility

		(1)	(2)	(3)	(4)
Log Retention Rate × Top 1		1.376***	1.508***	1.451***	1.404***
		(0.478)	(0.486)	(0.489)	(0.489)
Log Retention Rate × Top 1-5		0.926**	1.065**	1.004**	0.950**
		(0.449)	(0.455)	(0.458)	(0.457)
Log Retention Rate × Top 5-10		0.629	0.773*	0.713	0.654
		(0.449)	(0.455)	(0.457)	(0.456)
Log Retention Rate × Top 10-25		0.357	0.511	0.454	0.396
		(0.441)	(0.447)	(0.448)	(0.447)
Log Retention Rate × Below Top 25		0.0775	0.263	0.210	0.166
		(0.444)	(0.451)	(0.449)	(0.449)
Quality× Country FE		NO	YES	YES	YES
Quality × Country FE × Year		NO	NO	YES	YES
Quality×Country FE×Year×Field FE		NO	NO	NO	YES
Control: Top 5-10	Domestic elasticity	.02	.02	.02	.02
	s.e	(.005)	(.005)	(.005)	(.005)
	Foreign elasticity	.63	.62	.62	.63
	s.e	(.18)	(.18)	(.19)	(.19)
Control: Top 10-25	Domestic elasticity	.03	.02	.02	.02
	s.e	(.005)	(.005)	(.005)	(.005)
	Foreign elasticity	.85	.84	.83	.84
	s.e	(.18)	(.18)	(.18)	(.18)
Control: Below Top 25	Domestic elasticity	.03	.03	.03	.03
	s.e	(.005)	(.005)	(.006)	(.006)
	Foreign elasticity	1.09	1.05	1.04	1.04
	s.e	(.190)	(.196)	(.201)	(.203)
Observations		8644280	8616336	8616336	8616336

Notes: Multinomial logit regressions. Robust standard errors clustered at the inventor level in parentheses. See the explanatory notes to Table 4. All regressions contain the same covariates as Table 4, column 4, except that the country specific time trends, country GDP per capita and country patent stock have been replaced by country × year fixed effects. "Domestic elasticity" is the elasticity of top 1% superstar domestic inventors with respect to the top net-of-tax rate (one minus the top tax rate), while "Foreign elasticity" is the elasticity of top 1% superstar foreign inventors with respect to the top net-of-tax rate. They are computed, respectively, according to formulas (9) and (10) for three different control groups $g \in \{$top 5-10%, top 10-25%, below top 25%$\}$. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

rate by 10 percentage points from 60% (the average value of the top tax rate in the sample in year 2000) to 50%. Extrapolating from these elasticities and choosing the below top 25% group as a control group, the average country would be able to retain 1% more domestic top 1% superstar inventors and would be able to attract 26% more foreign superstar top 1% inventors.[46]

We can also compute the elasticities of domestic and foreign top 1% superstar inventors country by country, as in formulas (7) and (8). To avoid carrying too many numbers, let us pick the estimates obtained using the below top 25% of inventors as a control group. In Table 6, columns 1 and 2

[46] Of course, the percent difference here is due to the very different bases: there are many more domestic inventors in a country than foreign inventors.

show, respectively, the elasticity of domestic and foreign top 1% superstar inventors to the top net-of-tax rate. The U.S. and Japan stand out with their relatively low elasticities of both foreign and domestic top 1% superstar inventors, while superstar inventors from Great Britain and Canada are particularly elastic. France, Italy and Switzerland have moderately elastic domestic top 1% superstar inventors. Columns 3 and 4 compute the percent change in, respectively, domestic and foreign top 1% superstar inventors for a 10 percentage point top tax rate reduction from the actual level in each country. The differences in the percent changes are due to the different elasticities in the first two columns, but most importantly, to the very different bases on which they are computed (i.e., there are very different absolute numbers of domestic and foreign top 1% superstar inventors in the various countries). At the low end of the spectrum, the U.S. would retain 0.1% more domestic superstar inventors, Japan 0.2%, and Germany 1.4%. At the high end of the spectrum, Great Britain and Canada would retain, respectively, 8.1% and 6% more domestic superstar inventors. At the same time, the US would attract 18.3% more foreign superstar inventors, while France would attract 43.5% more superstar foreign inventors.[47]

Table 6: Migration Elasticities across Countries

Country	Domestic elasticity	Foreign elasticity	% change in domestic inventors	% change in foreign inventors
United States	0.003	0.97	0.1%	18.3%
Great Britain	0.37	1.23	8.1%	27.1%
Canada	0.31	1.23	6.0%	23.6%
Germany	0.05	1.21	1.4%	33.8%
France	0.12	1.23	4.4%	43.5%
Italy	0.13	1.23	3.0%	27.4%
Japan	0.01	1.23	0.2%	25.2%
Switzerland	0.18	1.23	4.1%	27.9%

Notes: Elasticities per country are calculated for the year 2000. The control group chosen is the below top 25% of inventors from Table 5. To compute the percent change in domestic inventors and the change in foreign inventors for each country, we consider a 10 percentage points decrease in taxes from their actual level in that country in year 2000. Similar changes can of course be computed for any other percentage point change in taxes.

Economic Gains from Lower Taxes: A back of the envelope calculation can reveal the yearly economic gain from lower taxes through the channel of inventor migration. Suppose that top retention rates in country c change by $d(1-\tau_{ct})$. We compute the economic value gained from attracting more domestic and foreign top 1% superstar inventors as:

$$dV_{ct} = \frac{d(1-\tau_{ct})}{(1-\tau_{ct})} \times (\varepsilon_d^c \times N_c^d + \varepsilon_f^c \times N_c^f) \times N_p \times V_p \tag{11}$$

where ε_d^c and ε_f^c are as defined in equations (7) and (8), N_c^d and N_c^f are the number of, respectively,

[47] Note that the heterogeneous elasticities of migration to tax rates across countries could be another reason for why some countries can tax more than others (Kleven, 2014).

domestic and foreign top 1% superstar inventors who live in country c, N_p is the average number of patents per year of top 1% superstar inventors (2.7 in our sample), and V_p is the average value per patent.

To assign an average value per patent, we use estimates from the literature. Pakes (1985) finds that the average value of a patent is \$ 810,000 (in 1972 dollars), which represents \$ 2.7 million in 2000 dollars.[48] Naturally, this is likely to be a lower bound on the value of the patents from superstar top 1% inventors. Superstar inventors are by definition those who create the most valuable breakthrough patents. Indeed, while the median patent in our sample receives 8 citations, the most valuable top 1% patents receive on average 171 citations, which is 21 times higher. Hence, we can perform a second calculation, supposing that the value per patent of superstar top 1% inventors is close to the value of the top 1% most valuable patents (which could be approximated by 21 times the average value of \$ 2.7 million, i.e., \$ 57 million.[49]) The potential gains in economic value from top tax decreases are shown in Table 7 for a 5 percentage point and a 10 percentage point top tax decrease. Naturally, these are only the gains through the migration channel of top 1% inventors and there are of course many other economic gains or losses from reducing the top tax rate through other channels in the rest of the economy.[50] These effects could also be a lower bound of the economic value lost if there are positive spillovers from having superstar inventors locally (as documented in the papers cited in the Introduction and in Section 2.3). As we show in Section 6.1, superstar inventors also have more patent breadth and breadth of impact, above and beyond pure citations, which may imply that their patents are even more valuable.

The economic gains per country from decreasing top tax rates depend on the domestic and foreign elasticities (in Table 6), but also on the size of the inventor base. The U.S. has the largest economic gain, despite the small elasticities, while Italy has the smallest economic gains, which is a combination of a relatively low elasticity of domestic inventors and a relatively small inventor base. To a first-order we could imagine that these yearly gains would scale up for a tax decrease that lasts several years.

5.4 The Role of Companies

What is the role of employers in determining inventors' mobility? One might expect that role to be important. Indeed, large companies often recruit internationally. In the U.S., Qualcomm Inc. and Microsoft have ratio of foreign inventors to total inventors in the company of, respectively, 51% and 57%. In Switzerland, Alstom Technology Ltd. and Syngenta Participation AG both have immigration rates of 67%.[51] There are two related questions when talking about the role of the

[48] We use the GDP deflator provided by the Bureau of Economic Analysis to adjust the nominal numbers.

[49] Assuming that the value of patents scales approximately linearly in citations.

[50] Because this is clearly just a very small fraction of all the effects that lower top tax rates could have, these numbers are not meant to be compared to the lost tax revenue. They are just meant to illustrate the "dollar value" of the elasticities estimated.

[51] WIPO report available at `http://www.wipo.int/edocs/pubdocs/en/intproperty/941/wipo_pub_941_2013.pdf`.

Table 7: Yearly Economic Gains Across Countries (in million USD)

Country	Small Patent Value		Large Patent Value	
	5% points tax change	10% points tax change	5% points tax change	10% points tax change
United States	58.0	116.1	1,225.5	2,451.0
Great Britain	16.4	32.7	345.5	691.0
Canada	17.6	35.1	370.6	741.3
Germany	17.7	35.4	373.2	746.5
France	10.9	21.7	229.1	458.3
Italy	3.0	5.9	62.7	125.3
Japan	8.5	17.0	180.0	360.0
Switzerland	5.5	11.0	116.0	232.0

Notes: All numbers are in millions. Elasticities per country are calculated for the year 2000. The control group chosen is the below top 25% of inventors from Table 5. To compute the change in domestic inventors and the change in foreign inventors, we consider a 5 and 10 percentage points decrease in top tax rates from the actual tax rate in each country. The economic gain per year is computed as in formula (11). Columns 1 and 2 assume a patent value per superstar top 1% inventor equal to the average patent value of $ 2.7 million. Columns 3 and 4 assume a large patent value equal to $ 57 million. These values are obtained using patent value estimates from Pakes (1985) and our own calculations provided in the text. These effects could be underestimating the economic value lost if there are positive spillovers from having superstar inventors locally (as documented in the papers cited in the Introduction and in Section 2.3).

employer. The first is whether employer characteristics, in particular whether it is a multinational company and how concentrated it is in some countries, matters. The second is whether the employer contributes to the decision of where to locate and, as a result, the estimated elasticity is a mix of employee and firm responses.

Regarding the latter question, the top personal income tax rate affects the surplus available from a firm-employee match. Depending on the bargaining setup between firms and employees, the firm should internalize to varying degrees the level of the personal income tax rate. In the limit, if the firm had to pay full compensating differentials for higher top tax rates to star workers, it should perfectly internalize personal income tax rates. At the same time, the company might bring to the table other considerations for relocating the worker, which are orthogonal to the personal income tax rate. In this latter case, the observed sensitivity to the top personal income tax rates should be lowered relative to a case where workers and firms decide unitarily. To sum up, any observed response to personal income taxes may be driven by the employee or by the firm, depending on the bargaining and wage setting process.

Regarding the role of the type of employer, in theory, the effects on international mobility of working for a multinational company are ambiguous. On the one hand, it might make an international move easier for the worker either directly within the company, or by giving him access to a wider international network built over the course of his career in a multinational. Of course, workers more likely to take advantage of tax differentials in the future could also simply self-select into a multinational with the expectation to move in response to tax rates. On the other hand, multinationals may be more able to relocate their workers internationally for strategic corporate reasons, rather than because of personal income tax rates, which would tend to dampen the estimated elasticity to top personal tax rates. The role of companies for inventor mobility is studied in Table 8.

Table 8: The Role of Companies for Inventor Mobility

		(1)	(2)
Log Retention Rate × Top 1		1.400***	0.980*
		(0.500)	(0.537)
Log Retention Rate × Top 1-5		0.868*	0.548
		(0.473)	(0.493)
Log Retention Rate × Top 5-10		0.514	0.199
		(0.473)	(0.491)
Log Retention Rate × Top 10-25		0.181	-0.0974
		(0.468)	(0.481)
Log Retention Rate × Below Top 25		-0.254	-0.560
		(0.472)	(0.485)
Log Retention Rate × Not Multinational		-0.216*	
		(0.129)	
Log Retention Rate × Activity abroad			-1.470***
			(0.137)
Quality× Country FE		YES	YES
Quality × Country FE × Year		YES	YES
Quality×Country FE×Year×Field FE		YES	YES
Control: Top 5-10	Domestic elasticity	.018	.011
	s.e	(.0045)	(.0047)
	Foreign elasticity	.809	.420
	s.e	(.201)	(.154)
Control: Top 10-25	Domestic elasticity	.024	.016
	s.e	(.0045)	(.0046)
	Foreign elasticity	1.113	.579
	s.e	(.197)	(.151)
Control: Below Top 25	Domestic elasticity	.034	.027
	s.e	(.0047)	(.0049)
	Foreign elasticity	1.511	.828
	s.e	(.211)	(.159)
Observations		7059856	6168504

Notes: Multinomial logit regressions. Robust standard errors clustered at the inventor level in parentheses. See the explanatory notes to Table 5. All regressions contain the same covariates as Table 5, column 4. Column 1 adds an interaction term of the log retention rate with whether the inventor was employed by a non-multinational company in the last period observed. Employees from non multinationals appear significantly less sensitive to the tax rate, which can indicate that they are less able to move internationally to take advantage of lower taxes. Column 2 adds as a control the share of innovative activity of the inventor's company that takes place in the destination country and an interaction of the activity share with the top retention rate. Inventors are significantly less sensitive to the top tax rate in any destination country if their company has a large share of its innovative activity in that country. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

Column 1 of Table 8 reports the results from the benchmark specification, adding the interaction of the top retention rate with a dummy equal to 1 if the inventor did *not* work for a multinational company in the previous period during which he was observed in the sample. Inventors who are not in multinationals are significantly less sensitive to the top tax rate. It seems, then, that working for a multinational company facilitates taking advantage of tax differentials.

Next, we compute the share of innovative activity of the inventor's company that takes place in each potential destination country, defined as the fraction of yearly patents of the company assigned to inventors from that country. Column 2 shows that if the company of the inventor has a higher share of its innovative activities in some destination country, the inventor is less sensitive to the retention rate in that destination country. The elasticity for those inventors whose company has no share of activity in the foreign country is very large, but declines very rapidly in the share of activity abroad.

This does not give us much indication as to who (the employee or the firm) ultimately takes the decision where to locate the worker. Indeed, a higher activity share abroad might mean that the idiosyncratic preference μ of the inventor is large for that destination country, due to career concerns, and this difference is sufficiently large to outweigh tax differentials. But it may instead be the company that wants to relocate inventors to high activity places for strategic corporate reasons unrelated to the personal income tax. Both scenarios would generate the reduced elasticity to tax rates observed in column 2.

While companies may or may not be relocating workers due to personal tax considerations, we can at least provide some evidence that they do not relocate inventors in our sample based on corporate taxation. Appendix Table A14 reports the results from regressions identical to those in Table 4, to which we add the corporate tax as an explanatory variable. The corporate tax rate has no explanatory power for inventors' location decisions.

Overall, even though we cannot ultimately detect who makes the decision to move, we showed that employer characteristics matter for the elasticity of migration to top tax rates.

6 Robustness to Alternative Specifications and Extensions

This section provides an extensive series of robustness checks on our benchmark results and several extensions.

6.1 Alternative inventor quality measures

Our measure of inventor quality and the corresponding ranking of inventors is central to our analysis. Therefore, we explore here how the results change if we consider different quality measures.

Static versus dynamic measures, wage formation and employer expectations: Recall that our benchmark quality measure is dynamic citations-weighted patents. This means that an inventor could be classified as a superstar today (if he is ranked in the top of the regional quality distribution in that year), but may not have been a superstar yesterday and may also no longer be a superstar tomorrow. I.e., dynamic superstar rankings can change over time and it is assumed that only the current quality ranking matters for income. This corresponds to a situation in which the employer has no foresight at all about the inventor's future potential or there are no job search and matching frictions that would prevent firms from firing a worker and hiring a better one in a future period in which information about the inventor's quality is revealed. Symmetrically, it also implies that, conditional on the current value of the quality measure and rank, past rankings do not matter.

Of course, the wage setting process is a complicated one and might depend both on past performance and future expectations, which provide the employer with additional information about the inventor's productivity not necessarily reflected in his current patents to date. To take a concrete example from academia, a university might hire a promising young Ph.D. graduate without any publication or patent record, based on his revealed skills during conversations and seminars. This new hire may have a high marginal product as measured by his impact on graduate students and colleagues that is not yet reflected in his publication record. Alternatively, a senior professor with several very high impact publications, but who is no longer producing at a level that would classify him in the current top 1% of researchers is probably still highly paid based on his past record.

If these situations apply in practice, a more appropriate component of the income of an inventor would be based on a lifetime, static quality. In particular, we can redefine an inventor as being in the top 1% of inventors if, over the course of his life in the sample, he has ever been or will ever be in the top 1%. The results for this alternative quality measure are in column 1 of Table 9. The elasticity measures are very close to the benchmark (column 4 in Table 5). In reality, wages are probably set to some extent based on current marginal product and to some extent based on potential future information and on realized past information, especially if there are search and matching frictions that make the employer want to lock in promising employees. It is hence reassuring that even a fully static measure still maintains our main result.

We also consider the case in which an inventor is defined as a superstar if he has ever been in the top 1% *in the past* (even if he currently no longer is in the top 1%). This corresponds to the case in which an inventor is always rewarded for past success, even independent of his current performance, but is not rewarded in anticipation of future success, or there is no advance, lead information about the inventor. The results are almost perfectly identical to the static measure in column 1 of Table 9 and hence not reported to save on space.

Patent quality versus quantity measures: As described in section 2.3, there are several possible quality measures based on patents and citations that put different emphasis on the quality versus the quantity of patents. The correlations of these various measures, although significantly positive at the 5% level, are not always large. For instance, the correlation between measure $q1$ (citations-weighted patents) and $q2$ (number of patents) is 0.70, the correlation between $q1$ and $q3$

Table 9: Robustness Checks

		Alternative quality Measures				Imputing location
		(1)	(2)	(3)	(4)	(5)
Log Retention Rate × Top 1		1.363***	0.326	2.559***	1.696***	1.532***
		(0.475)	(0.518)	(0.498)	(0.489)	(0.434)
Log Retention Rate × Top 1-5		1.131**	0.489	2.097***	1.313***	1.153***
		(0.455)	(0.454)	(0.456)	(0.455)	(0.402)
Log Retention Rate × Top 5-10		0.700	0.476	1.398***	0.738	0.940**
		(0.460)	(0.445)	(0.452)	(0.454)	(0.397)
Log Retention Rate × Top 10-25		0.415	0.594	0.738*	0.328	0.747*
		(0.450)	(0.441)	(0.446)	(0.448)	(0.390)
Log Retention Rate × Below Top 25		-0.0895	1.441***	0.183	0.123	0.810**
		(0.512)	(0.444)	(0.445)	(0.448)	(0.389)
Quality× Country FE		YES	YES	YES	YES	YES
Quality × Country FE × Year		YES	YES	YES	YES	YES
Quality×Country FE×Year×Field FE		YES	YES	YES	YES	YES
Control: Top 5-10	Domestic elasticity	.013	-.001	.011	.021	.015
	s.e	(.004)	(.0042)	(.0024)	(.0055)	(.0054)
	Foreign elasticity	.552	-.134	1.118	.84	.506
	s.e	(.167)	(.249)	(.262)	(.224)	(.188)
Control: Top 10-25	Domestic elasticity	.019	-.003	.015	.028	.019
	s.e	(.0038)	(.0042)	(.0024)	(.0052)	(.0053)
	Foreign elasticity	.788	-.241	1.752	1.199	.672
	s.e	(.162)	(.249)	(.254)	(.212)	(.182)
Control: Below Top 25	Domestic elasticity	.029	-.018	.021	.034	.018
	s.e	(.0061)	(.0045)	(.0023)	(.0051)	(.0054)
	Foreign elasticity	1.209	-1.008	2.286	1.378	.617
	s.e	(.258)	(.27)	(.248)	(.207)	(.185)
Observations		8616336	8616336	8616336	8616336	17173520

Notes: Multinomial logit regressions. Robust standard errors clustered at the inventor level in parentheses. See the notes to Table 5. All regressions contain the same covariates as Table 5, column 4. Column 1 uses the lifetime measure of inventor ranking ("has ever been or will ever be in the top 1%," etc..). The next three columns use three alternative quality measures to rank inventors into the top 1%, top 1-5%, etc.. Column 2 uses the patent count to date (measure $q2$ as defined by formula (2)). Column 3 uses the inventor's average citations per patent to date (measure $q3$ as defined in (3)). Column 4 uses the max citations per patent of an inventor to date (measure $q4$ as defined in (4)). Column 5 is based on a balanced panel, where an inventor's location for years in between patents is imputed according to the procedure described in the text. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

(average citations per patent) is 0.32 and the correlation between $q2$ and $q3$ is only 0.05 (see the full correlation matrix in Appendix Table A1).

We may be worried that our benchmark quality measure (citations-weighted patent, $q1$) is correlated with the likelihood of observing an inventor in the sample, and, consequently, with the likelihood of observing a move. Therefore, we provide three more tests by using patent counts ($q2$), average citations per patent ($q3$), and max citations per patent ($q4$) as alternative measures to show that it is really the quality of an inventor that matters, rather than the pure quantity of patents.

Column 2 of Table 9 reassures us that the results are not artificially driven by observing people with more patents more frequently. Using simple patent counts, the coefficient on the top retention rate becomes much smaller and insignificant for top 1% inventors, while positive and significant for the below top 25% inventors. On the other hand, using purely quality driven measures such as average citations per patent (in column 3) or max citations per patent (in column 4), maintains or even strengthens our benchmark results. Hence, the relation is not mechanically driven by frequency of observation.

Patent breadth and breadth of impact measures: Another way to assess an inventor's impact is through the breadth of his innovations, and the extent of his new and original ideas. We hence construct two additional quality measures. The first measure, which we call "breadth of impact" considers the number of technology classes that end up building on an inventor's patents. Formally, we take the set of patents of an inventor until time $t-1$ and count the number of technological classes which contain patents that ever cite those patents. The second measure, called patent breadth, is the dynamic claims-adjusted patent stock, i.e., the number of claims on all patents received by the inventor by time $t-1$ (constructed exactly as our benchmark measure $q1$, but using claims instead of citations). It is common in the literature to use patent claims to proxy for patent breadth (Lerner, 1994; Lanjouw and Schankerman, 2004). In Appendix Table A10, we show that, as expected, superstar inventors according to our benchmark definition, also have significantly more patent breadth and breadth of impact. The results using the breadth of impact and the patent breadth as quality measures are, respectively, in columns 1 and 2 of Table 10.

6.2 Accounting for potential selection

One problem about using the patent data to track inventors' locations is that we only observe inventors in years in which they patent, but not in the years between consecutive patents. We perform two analyses to address this problem. First, we make sure that the results do not change much when we impute the inventors' locations for missing years. Second, we estimate a selection model.

Imputing location for missing years: To impute observations for years in which inventors do not patent, we use the following imputation algorithm: if the inventor is seen in country A in year X and in country B in year $Y > X$, we assume that he has been in country A until year $X + (Y - X)/2$ and in country B thereafter and until year Y. We do not impute years before the first or after the last patent. Column 5 of Table 9 shows that the elasticities are slightly smaller, but still very similar when we use the imputed data.

Binary Selection Model: To check for selection based on the number of patents, we also use a formal selection model. To simplify the computational burden, we only study the mobility between the U.S. and Canada. The U.S.-Canada corridor is very large and among inventors who migrate within this corridor, very few also migrate to another country. Among 863,406 U.S.-born inventors,

Table 10: Breadth of Impact and Patent Breadth

		(1)	(2)
Log Retention Rate × Top 1		1.335***	1.113**
		(0.517)	(0.500)
Log Retention Rate × Top 1-5		1.180**	0.856*
		(0.488)	(0.462)
Log Retention Rate × Top 5-10		1.029**	0.598
		(0.478)	(0.450)
Log Retention Rate × Top 10-25		0.734	0.437
		(0.474)	(0.445)
Log Retention Rate × Below Top 25		0.624	0.179
		(0.471)	(0.444)
Quality× Country FE		YES	YES
Quality × Country FE × Year		YES	YES
Quality×Country FE×Year×Field FE		YES	YES
Control: Top 5-10	Domestic elasticity	.007	.015
	s.e	(.0047)	(.0068)
	Foreign elasticity	.269	.429
	s.e	(.203)	(.205)
Control: Top 10-25	Domestic elasticity	.012	.018
	s.e	(.0047)	(.0068)
	Foreign elasticity	.527	.563
	s.e	(.2)	(.204)
Control: Below Top 25	Domestic elasticity	.015	.026
	s.e	(.005)	(.0074)
	Foreign elasticity	.623	.777
	s.e	(.22)	(.22)
Observations		8616336	8616336

Notes: Multinomial logit regressions. Robust standard errors clustered at the inventor level in parentheses. See the notes to Table 5. All regressions contain the same covariates as Table 5, column 4. Column 1 uses the breadth of impact quality measure described in the text. Column 2 uses the patent breadth quality measure. $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$

640 are observed only in Canada later, while 35 are observed both in Canada and some other country. Among 44,321 Canada-born inventors, 946 are observed only in the U.S. later and 18 are observed both in the U.S. and some third country.

First, we estimate a simple probit model for the inventors born and observed only in Canada or the U.S., where the dependent variable takes the value 1 if an inventor locates in the U.S. The dependent variables are analogous to the ones described in Section 5.1, with some small modifications. Again, we include the same set of controls i)-iii), and we consider progressively more detailed specifications for the counterfactual wages, by adding quality measures, quality-country specific trends and quality-country-field specific trends. The retention rate is the U.S. top retention rate, interacted with inventor rankings (top 1%, top 1-5%, ..).

Next, we perform a formal Heckman selection model with exactly the same controls and compare its results with those from the simple probit model without selection. We exploit a reform enacted in 1995 and introduced by the "Patent Term and Publication Reform Act of 1994" to synchronize patent terms with requirements of the Uruguay Round Agreements Act. The main changes were that the patent term of 17 years (counted from the patent grant year) was changed to 20 years from the patent's earliest application year. Given that on average the patent grant period is less than 3 years, this is typically an effective increase in the patent term. Indeed, the average lag between the application and the grant was 2 years, which effectively implies that the reform has increased the average duration of patent protection upon patent grant by 1 year to 18 years (=20 years − 2 years). Appendix Table A9 provides the average gap between the application and grant years on average, as well as for inventors of different qualities and industries. This reform increased the likelihood of patenting, and, hence, of observing inventors in the data (see Table 11 below), but should not have affected location choices. Hence, we use a dummy for post-1994 in the first stage of our selection model.

Table 11 shows the results from the simple probit model and the corresponding selection model. It confirms that the effects of top tax rates remain very similar after controlling for selection.

Table 11: Heckman Selection model on Canada-U.S

	(1) Probit	(2) Selection
US log retention rate × Top 1	1.302***	1.303***
	(0.283)	(0.283)
US log retention rate × Top 1 - 5	0.281	0.280
	(0.201)	(0.201)
US log retention rate × Top 5 - 10	0.202	0.199
	(0.156)	(0.156)
US log retention rate × Top 10 - 25	0.171	0.168
	(0.121)	(0.122)
US log retention rate × Below top 25	-0.0314	-0.0359
	(0.120)	(0.120)
First stage		
Post reform (1994) dummy		0.0882***
		(0.0288)
Observations	568749	1160136

Notes: Estimation on a sample limited to the United States and Canada. Probit regression in column (1) with dependent variable equal to 1 if the inventor locates in the United States. Heckman selection estimation in column (2), using a dummy for post-1994 as an instrument. Robust t-statistics clustered at the inventor level in parentheses. All columns contain the same covariates as column 3 of Table 5. The probit and selection models yield very similar coefficients on the interaction of top 1% inventors and the U.S. top retention rate. The first stage of the reform is highly significant. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

To sum up, the results using imputed data showed that imputing missing years leaves the results highly significant (Table 9, column 5). The results from section 6.1 using patent count as a quality measure showed that the relation between patenting and moving was not mechanical. Combined with the formal selection model results, this tends to show that selection based on patenting is likely not the driver of our results.

Sample selection: One might be worried that the choice of patenting in the USPTO is not innocuous for foreign inventors. First, it could be that inventors who plan to migrate to the U.S. tend to patent more heavily with the USPTO. This would tend to make the observed sample more mobile on average than a fictitious sample in which inventors are randomly assigned to patent offices. On the other hand, however, conditional on appearing in the sample, those who had planned to migrate to the U.S. (and, accordingly, had decided to patent with the USPTO) may not be driven at all by tax considerations, which would tend to reduce the elasticity to taxes. To address this issue we first repeat the analysis on inventors who patent with the EPO. Many more inventors in the EPO data are Europeans. These results are presented in Section 7 and confirm our previous results. Second, we drop all movers to the U.S. from both the USPTO and the EPO samples. These results are reported in Appendix Tables A15 and A17. These tables show that the effect of the top retention rate is still very significant on the superstar top 1% inventors.[52]

6.3 Long-term Mobility

The mobility of inventors has different economic implications if it is short-term versus long-term. In this section, we consider the effects of taxation exclusively on long-term mobility. We define a long-term move as a move that is never followed by a move back to the origin country during the time the inventor is in the sample. Recall that the average time in the benchmark sample for an inventor is 12 years. On the one hand, long-term mobility may be much less sensitive to contemporaneous tax rates, since people should normally re-optimize only once future tax changes happen. On the other hand, long-term mobility could be even more strongly correlated with tax changes if tax changes are highly persistent and people are aware of this fact, or if people overestimate the persistence of tax changes and are then faced with moving costs that prevent them from moving back during the average period of 12 years that we observe them. The results in column 1 of Table 12 show that domestic inventors' long-term mobility is slightly less sensitive to tax rates. For foreign inventors, it is slightly more sensitive to tax rates. When it comes to long-term moves, it does not seem to matter significantly whether the inventor works for a multinational (column 2). It still seems to matter how concentrated the research activity of the employer is (column 3), although the effects are weakened. This may indicate that for long-term mobility career concerns continue to matter.

[52] Note that for the USPTO sample, overall the estimated coefficients are larger, but the elasticities for domestic superstar inventors are a bit smaller and the elasticities of foreign superstar inventors are a bit larger. This is due to the lower overall number of moves once we drop moves to the U.S. (see our discussion of how the elasticities are derived from the multinomial logit as in (9) and (10) and related to the probability of staying in one's home country). For the EPO data described in Section 7, there is barely any change at all in the estimated elasticities of domestic superstar inventors from dropping movers to the U.S., and an increase in the elasticity of foreign superstar inventors.

Table 12: Long-term mobility

		(1)	(2)	(3)
Log Retention Rate × Top 1		2.600***	2.285***	2.168***
		(0.781)	(0.798)	(0.817)
Log Retention Rate × Top 1-5		2.273***	1.677**	1.687**
		(0.749)	(0.772)	(0.742)
Log Retention Rate × Top 5-10		1.705**	1.265*	1.296*
		(0.749)	(0.763)	(0.728)
Log Retention Rate × Top 10-25		1.211	0.997	1.009
		(0.742)	(0.760)	(0.721)
Log Retention Rate × Below Top 25		1.067	0.549	0.563
		(0.741)	(0.767)	(0.726)
Log Retention Rate × Not Multinational			-0.148	
			(0.165)	
Log Retention Rate × Activity abroad				-1.636***
				(0.190)
Quality× Country FE		YES	YES	YES
Quality × Country FE × Year		YES	YES	YES
Quality×Country FE×Year×Field FE		YES	YES	YES
Control: Top 5-10	Domestic elasticity	.012	.011	.007
	s.e	(.0042)	(.0032)	(.0038)
	Foreign elasticity	.777	.954	.458
	s.e	(.27)	(.279)	(.218)
Control: Top 10-25	Domestic elasticity	.021	.013	.009
	s.e	(.0041)	(.0031)	(.0037)
	Foreign elasticity	1.206	1.204	.61
	s.e	(.254)	(.271)	(.215)
Control: Below Top 25	Domestic elasticity	.022	.019	.014
	s.e	(.0041)	(.0033)	(.0038)
	Foreign elasticity	1.331	1.624	.846
	s.e	(.259)	(.288)	(.223)
Observations		8413144	6880848	6011504

Notes: Multinomial logit regressions. Robust standard errors clustered at the inventor level in parentheses. See the notes to Table 5. All regressions contain the same covariates as Table 5, column 4. The sample is restricted to only include long-term movers, defined as inventors who move without coming back during their time in the sample. Column 2 adds an interaction of the log retention rate with not having being the employee of a multinational firm in the previous period. Column 3 adds as a control the share of activity of the inventor's company in the destination country and an interaction of the activity share with the top retention rate. $^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$

7 European Patent Office Inventors

All our micro results until now used the DID, derived from the USPTO data. This data is naturally heavily dominated by U.S. inventors. Therefore, it is very valuable to test whether the sensitivity of superstar inventors to top tax rates also holds in the European Patent Office (hereafter, EPO) data. The disambiguation of the EPO data has only recently been done and is the subject of on-going research efforts (see Coffano and Tarasconi (2014), Breschi, Lissoni, and Tarasconi (2014)).

In this data, the representation of the same 8 countries we used for our benchmark analysis is as follows: Canada accounts for 1.3%, Switzerland for 3.3%, Germany for 23.7%, France for 7.7%, Great Britain for 6.2%, Italy for 3.8%, Japan for 16.4% and the U.S. for 27.5%. Hence, the biggest difference is the reduced share of the U.S. in the data. The construction of all variables and the sample is the same as for the benchmark analysis. For consistency, we also consider the exact same 8 countries. Appendix Table A11 provides some summary statistics for the European patent office data. The sample period here is 1977-2007.

Table 13 provides the counterpart of the benchmark results in Table 5 and the extended results in Table 9 for the EPO data.

TABLE 13: INVENTOR MOBILITY IN THE EUROPEAN PATENT OFFICE DATA

		Benchmark	Alternative quality Measures			
		(1)	(2)	(3)	(4)	(5)
Log Retention Rate × Top 1		0.978***	1.145***	2.257***	1.737***	0.188
		(0.320)	(0.285)	(0.358)	(0.313)	(0.317)
Log Retention Rate × Top 1-5		0.954***	0.944***	1.626***	1.054***	0.492**
		(0.240)	(0.232)	(0.261)	(0.253)	(0.229)
Log Retention Rate × Top 5-10		0.840***	0.741***	1.352***	0.956***	0.647***
		(0.230)	(0.243)	(0.247)	(0.243)	(0.219)
Log Retention Rate × Top 10-25		0.295	0.347	0.888***	0.445**	0.989***
		(0.224)	(0.233)	(0.234)	(0.226)	(0.213)
Log Retention Rate × Below Top 25		0.196	0.0189	0.00663	-0.0526	1.312***
		(0.264)	(0.593)	(0.228)	(0.248)	(0.256)
Quality× Country FE		YES	YES	YES	YES	YES
Quality × Country FE × Year		YES	YES	YES	YES	YES
Quality×Country FE×Year×Field FE		YES	YES	YES	YES	YES
Control: Top 5-10	Domestic elasticity	.001	.007	.004	.014	-.009
	s.e	(.0057)	(.0042)	(.0016)	(.0048)	(.0048)
	Foreign elasticity	.136	.395	.902	.765	-.450
	s.e	(.273)	(.253)	(.311)	(.259)	(.268)
Control: Top 10-25	Domestic elasticity	.014	.012	.007	.021	-.014
	s.e	(.0054)	(.0038)	(.0016)	(.0047)	(.0047)
	Foreign elasticity	.667	.780	1.366	1.269	-.782
	s.e	(.266)	(.235)	(.307)	(.247)	(.265)
Control: Below Top 25	Domestic elasticity	.02	.02	.012	.032	-.019
	s.e	(.0065)	(.0096)	(.0016)	(.0049)	(.0057)
	Foreign elasticity	.763	1.102	2.242	1.756	-1.099
	s.e	(.308)	(.583)	(.302)	(.26)	(.319)
Observations		8461393	8461393	8461393	8461393	8461393

Notes: Multinomial logit regressions on data from the EPO. Robust standard errors clustered at the inventor level in parentheses. See the explanatory footnote to Table 5. All columns contains the same covariates as Column 4 in Table 5. Column 1 corresponds exactly to column 4 from Table 5, but using the EPO data. Column 2 uses the static or lifetime measure of inventor ranking ("has ever been or will ever be in the top 1%," etc..). The next three columns use three alternative quality measures to rank inventors into the top 1%, top 1-5%, etc.. Column 3 uses the inventor's average citations per patent to date (measure $q3$ as defined in (3)). Column 4 uses the max citations per patent of an inventor to date (measure $q4$ as defined in (4)). Column 5 uses the patent count to date (measure $q2$ as defined by formula (2)). All these quality measures were described in detail in Sections 2.3 and 6.1. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$

Column 1 contains the results using our benchmark quality measure of citations-weighted patents ($q1$). We see the same pattern as in the DID with the superstar top 1% inventors very significantly sensitive to top tax rates. The elasticities of both domestic and foreign superstar top 1% inventors are somewhat smaller than in the DID. Indeed, in the EPO, even the top 5-10% inventors are significantly sensitive to top tax rates, probably because top tax brackets are in general lower in European countries (which are naturally more heavily represented in the EPO than in the DID). Using the top 10-25% and the below top 25% as control groups, yields, respectively, elasticities of domestic top 1% superstar inventors of 0.014 and 0.02 and elasticities of foreign top 1% superstar inventors of 0.67 and 0.76.

The next four columns consider the alternative quality measures already shown for the DID in Table 9. Column 2 shows the static or lifetime quality measure ("has ever been or will ever be in the top 1%," etc...) as described in Section 6.1. Column 3 shows the average citations per patent (measure $q3$), while column 4 shows the max citations per patent (measure $q4$). As was the case in the benchmark DID as well, ranking inventors according to their max citations per patent (measure $q4$) yields the strongest elasticities for domestic top 1% superstar inventors of, respectively 0.02 and 0.032 depending on the control group. Column 5 reports the results for the pure patent count, for which, as explained in Section 6.1, it is a very good sign that the effects of the top retention rate on superstar inventors are not significant.

In Online Appendix Table A17, we provide the aforementioned results with the EPO data, dropping all inventors who ever move to the U.S.. The effect of the top tax rate on superstar top 1% inventors remains strongly significant.

Overall, the results using the EPO data are highly consistent with those using the benchmark DID, which gives further confidence in our estimated effects of the top tax rate on superstar top 1% inventors.

8 Conclusion

In this paper, we consider the effects of taxation on the international mobility of inventors, who are economically highly valuable agents and key drivers of economic growth. We put particular emphasis on superstar inventors, those with the most and most valuable inventions. We use disambiguated inventor data, based on USPTO and EPO data, to track the international location of inventors over time, and combine it with effective top marginal tax rate data. We exploit variations in the top tax rate across time and countries, as well as its differential impact on inventors at different points in the quality distribution.

We first provided stylized macroeconomic facts that highlight the responses of superstar inventors to top taxes. We then presented quasi-experimental evidence from three country case studies that exploited large changes in migration policies or tax policy. We estimate a multinomial location model and find that the baseline elasticity of the number of superstar domestic inventors to the top net-of-tax rate is small (around 0.03 in our preferred specification). This translates on average into an increase of 1% in domestic superstar top 1% inventors at home for a 10 percentage points decrease in top tax rates from a level of 60%. The elasticity of the number of superstar top 1% foreign inventors to the net-of-tax rate is much higher, around 1, which translates into a 26% increase in foreign superstar top 1% inventors for the same 10 percentage points decrease in top tax rates. Inventors who have worked for multinationals in the previous period are more likely to take advantage of tax differentials, possibly because working for a multinational makes a move abroad easier and grants the inventor international exposure. On the other hand, inventors whose company has a research activity that is highly concentrated in a given country are less sensitive to tax differentials in that country, presumably because career concerns (being located where the company's main research activity is) outweigh tax considerations. We also find evidence for sorting effects by ability and negative spillovers from high quality to low quality inventors.

We then perform extensive robustness checks on the measures of earnings and quality used, the length of the migration spells studied, and potential selection based on patenting behavior. In these extensions and robustness checks, our results persist. We also reproduce the analysis on a new disambiguated inventor dataset based on European Patent Office patents and find baseline elasticities to the net-of-tax rate of 0.02 for domestic top 1% superstar inventors and of 0.76 for foreign top 1% superstar inventors.

These results suggest that, if the economic contribution of these key agents is very important, their migratory responses to tax policy might represent a cost to tax progressivity. Our estimates could fruitfully be used to calibrate the models of optimal taxation in the presence of migration cited in the Introduction. An additional relevant consideration is that inventors may have strong spillover effects on their geographically close peers, making it even more important to attract and retain them domestically.

Because inventors are key determinants of economic growth, this paper speaks to the relation between taxation and growth. An interesting direction for future research would be to include the migration margin of inventors, together with their externalities, into a structural economic growth model with taxation.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller, "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, 2010, *105* (490), 493–505.

Abrams, David, Ufuk Akcigit, and Jillian Popadak, "Patent Value and Citations: Creative Destruction or Strategic Disruption?," Working Paper 19647, National Bureau of Economic Research 2013.

Bakija, Jon and Eugene Steuerle, "Individual Income Tax since 1948," *National Tax Journal*, 1991, *44* (4), 451–475.

_ **and Joel Slemrod**, "Do the Rich Flee from High State Taxes? Evidence from Federal Estate Tax Returns," Working Paper 10645, National Bureau of Economic Research July 2004.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen, "The Lifecycle of Inventors," Working Paper in progress, Harvard University 2014.

Bessen, James, "The value of U.S. patents by owner and patent characteristics," *Research Policy*, 2008, *37*, 932–945.

Breschi, S., F. Lissoni, and G. Tarasconi, "Inventor Data for Research on Migration and Innovation: A Survey and a Pilot," Economic Research Working Paper 17, World Intellectual Property Organization 2014.

Chesbrough, Henry, "Emerging Secondary Markets for Intellectual Property: US and Japan Comparisons," Technical Report, National Center for Industrial Property Information and Training 2006.

Chiquiar, Daniel and Gordon H. Hanson, "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States," *Journal of Political Economy*, 2005, *113* (2), 239–281.

Coffano, Monica and Gianluca Tarasconi, "CRIOS - Patstat Database: Sources, Contents and Access Rules," Technical Report 1, Center for Research on Innovation, Organization and Strategy 2014.

Cosar, A. Kerem, Nezih Guner, and James Tybout, "Firm Dynamics, Job Turnover, and Wage Distributions in an Open Economy," Working Paper 16326, National Bureau of Economic Research September 2010.

Eissa, Nada, "Taxation and Labor Supply of Married Women: The Tax Reform Act of 1986 As a Natural Experiment," Working Paper 5023, NBER 1995.

_ **and Hilary Hoynes**, "The Hours of Work Response of Married Couples: Taxes and the Earned Income Tax Credit," in Jonas Agell and Peter Birch Sorensen, eds., *Tax Policy and Labor Market Performance*, MIT Press, 2006.

Foley, Fritz C. and William R. Kerr, "Ethnic Innovation and U.S. Multinational Firm Activity," *Management Science*, 2013, *59* (7), 1529–1544.

Gambardella, A. and Giuri, P. and Harhoff, D. and Hoisl, K. and Mariani, M. and Nagaoka, S. and Torrisi, S., "Invention Processes and Economic Uses of Patents: Evidence from the PatVal 2 Survey," Technical Report 2014.

Giuri, Paola, Myriam Mariani, Stefano Brusoni, Gustavo Crespi, Dominique Francoz, Alfonso Gambardella, Walter Garcia-Fontes, Aldo Geuna, Raul Gonzales, Dietmar Harhoff, Karin Hoisl, Christian Le Bash Alessandra Luzzi, Laura Magazzini, Lionel Nesta, nder Nomaler, Neus Palomeras, Pari Patel, Marzia Romanelli, and Bart Verspagen, "Inventors and Invention Processes in Europe: Results from the PatVal-EU Survey," *Research Policy*, 2007, *36* (8), 1107–1127.

Gravelle, Jane G. and Thomas L. Hungerford, "The Challenge of Individual Income Tax Reform: An Economic Analysis of Tax Base Broadening," CRS Report for Congress 7-5700, Congressional Research Service 2012.

Grogger, Jeffrey and Gordon H. Hanson, "Income maximization and the selection and sorting of international migrants," *Journal of Development Economics*, 2011, *95* (1), 42–57.

_ **and** _ , "Attracting Talent: Location Choices of Foreign Born PhDs in the US," *Journal of Labor Economics*, 2015, pp. 493–505.

Guner, Nezih, Martin Lopez-Daneri, and Gustavo Ventura, "Heterogeneity and Government Revenues: Higher Taxes at the Top?," mimeo, Arizona State 2014.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools," Working Paper 8498, National Bureau of Economic Research 2001.

Harhoff, Dietmar and Karin Hoisl, "Institutionalized Incentives for Ingenuity: Patent Value and the German Employees Inventions Act," *Research Policy*, 2007, *36*, 1143–1162.

_ , F. Narin, F. Scherer, and K. Vopel, "Citation Frequency and the Value of Patented Inventions," *Review of Economics and Statistics*, 1999, *81* (3), 511–515.

Hoynes, Hilary W. and Diane Whitmore Schanzenbach, "Consumption Responses to In-Kind Transfers: Evidence from the Introduction of the Food Stamp Program," *American Economic Journal: Applied Economics*, 2009, *1* (4), 109–139.

_ and _ , "Work incentives and the Food Stamp Program," *Journal of Public Economics*, 2012, *96*, 151–162.

Kerr, William R., "The Ethnic Composition of U.S. Inventors," Working Papers 08-006, Harvard Business School 2008.

_ , "U.S. High-Skilled Immigration, Innovation, and Entrepreneurship: Empirical Approaches and Evidence," Working Papers 14-017, Harvard Business School 2013.

Kleven, Henrik, "How Can Scandinavians Tax So Much?," *Journal of Economic Perspectives*, 2014, *28* (4), 77–98.

Kleven, Henrik Jacobsen, Camille Landais, and Emmanuel Saez, "Taxation and International Migration of Superstars: Evidence from the European Football Market," *American Economic Review*, 2013, *103* (5), 1892–1924.

_ , _ , _ , and Esben Schultz, "Migration and Wage Effects of Taxing Top Earners: Evidence from the Foreigners Tax Scheme in Denmark," *The Quarterly Journal of Economics*, 2014, *129* (1), 333–378.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman, "Technological Innovation, Resource Allocation, and Growth," Working Paper 17769, National Bureau of Economic Research 2012.

Lai, Ronald, Alexander D'Amour, David M. Doolin, Guan-Cheng Li, Ye Sun, Vetle Torvik, Amy Yu, and Lee Fleming, "Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database," Working Paper, Harvard Dataverse Network 2012.

Lanjouw, Jean O. and Mark Schankerman, "Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators," *Economic Journal*, 2004, *114* (495), 441–465.

Lehmann, Etienne, Laurent Simula, and Alain Trannoy, "Tax Me if You Can! Optimal Nonlinear Income Tax between Competing Governments," *Quarterly Journal of Economics*, 2014, *129* (4).

Lerner, Josh, "The Importance of Patent Scope: An Empirical Analysis," *RAND Journal of Economics*, 1994, *25* (2), 319–333.

Liebig, Thomas, Patrick A. Puhani, and Alfonso Sousa-Poza, "Taxation and Internal MigrationEvidence from the Swiss Census Using Community-Level Variation in Income Tax Rates," *Journal of Regional Science*, 2007, *47* (4), 807–836.

Miguelez, Ernest, "Inventor Diasporas and the Internationalization of Technology," Discussion Paper 1425, CReAM 2013.

_ and Carsten Fink, "Measuring the International Mobility of Inventors: A New Database," Working Paper 8, WIPO 2013.

— **and Rosina Moreno**, "What Attracts Knowledge Workers? The Role of Space and Social Networks," *Journal of Regional Science*, 2014, *54* (1), 33–60.

Mirrlees, James, "Migration and Optimal Income Taxes," *Journal of Public Economics*, 1982, *18*, 319–341.

Moretti, Enrico and Daniel Wilson, "State Incentives for Innovation, Star Scientists and Jobs: Evidence from Biotech," *The Journal of Urban Economics, forthcoming*, 2014.

— **and** — , "The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists," Working Paper 21120, National Bureau of Economic Research 2015.

Moser, Petra, Joerg Ohmstedt, and Paul Rhode, "Patent Citations and the Size of Innovation ? Evidence from Hybrid Corn," Working Paper, Stanford University 2012.

NSF, "NSF SESTAT, National Survey of College Graduates (NSCG)," 2003.

Pakes, Ariel, "On Patents, R&D, and the Stock Market Rate of Return," *Journal of Political Economy*, April 1985, *93* (2), 390–409.

— , "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, 1986, *54* (4), 755–784.

— **and Mark Schankerman**, "Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period," *Economic Journal*, 1986, *96* (384), 1052–1076.

Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva, "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities," *American Economic Journal: Economic Policy*, 2014, *6* (1), 230–71.

Ryder, Rodney D and Ashwin Madhavan, *Intellectual Property and Business: The Power of Intangible Assets*, SAGE Publications, 2014.

Simula, Laurent and Alain Trannoy, "Optimal Income Tax under the Threat of Migration by Top-Income Earners," *Journal of Public Economics*, 2010, *94* (2), 163–173.

Toivanen, Otto and Lotta Väänänen, "Returns to Inventors," *The Review of Economics and Statistics*, 2012, *94* (4), 1173–1190.

Trajtenberg, Manuel, "A Penny for Your Quotes: Patent Citations and the Value of Innovations," *The RAND Journal of Economics*, 1990, *21* (1), 172–187.

Wilson, John D., "The Effect of Potential Emigration on the Optimal Linear Income Tax," *Journal of Public Economics*, December 1980, *14* (3), 339–353.

— , "Optimal Linear Income Taxation in the Presence of Emigration," *Journal of Public Economics*, August 1982, *18* (3), 363–379.

Young, Cristobal and Charles Varner, "Millionaire Migration and State Taxation of Top Incomes: Evidence from a Natural Experiment," *National Tax Journal*, 2011, *64* (2), 255–283.

Zucker, Lynne and Michael Darby, "Movement of Star Scientists and Engineers and High-Tech Firm Entry," *Annals of Economics and Statistics*, 2014, *115/116.*