# What are the effects of Corporate Tax on Innovation?


# Wilson Tai

Student ID: 21204700
Master of Science: Quantitative Economics
Supervisor: Professor Ron Davies

University College Dublin
UCD School of Economics



## Abstract

Accounting for innovation's inertial behaviours, this thesis examines the relationship between corporate tax and innovation output at a cross-country setting. Beginning with a cross-sectional dataset on patent applications, we build a micro-aggregate panel data for members of the European Union from 2000 to 2017. We find that an increase in the aggregate effective average corporate tax rate three years ago reduces innovation activity and the quality of innovation this year. In addition to corroborating our baseline results, our robustness checks also found that: 1) innovation activity is not effected by corporate tax until three years later; 2) corporate tax effects the quality of innovation up to the third year; and 3) the elasticity of innovation output relative to corporate tax is larger for the alternative micro-aggregate of solely star inventors than non-star inventors.

# Contents


**1 Introduction** **1**

**2 Literature Review** **3**

2.1 Innovation Measurements 3

2.2 Corporate Tax and Innovation 5

**3 Data** **7**

3.1 Patent Data 7

3.2 Country Data 11

3.3 Summary Statistics 12

3.4 Econometric Methodology 14

**4 Results** **15**

4.1 Baseline Results 15

4.2 Robustness Checks 17

4.2.1 Alternative Dynamics 17

4.2.2 Alternative Aggregation: Star Inventors 20

4.2.3 Alternative Aggregation: Technology Classes 22

4.2.4 Truncation Bias 25

**5 Conclusion** **26**

**6 References** **29**

**7 Appendix** **33**

7.1 Appendix A: Cooperative Patent Classification Explanation 33

7.2 Appendix B: Country-CPC Weight Example 34

# List of Figures


| | | |
|---|---|---|
| 1 | Forward Citations and Truncation Bias | 10 |
| 2 | Effective Average Tax Rate in 2000 and 2017 | 13 |


# List of Tables


| | | |
|---|---|---|
| 1 | Summary Statistics | 12 |
| 2 | Baseline PPML Regression | 15 |
| 3 | Alternative Dynamics | 18 |
| 4 | Alternative Aggregation: Star and Non-Star Inventors | 21 |
| 5 | Cross-CPC Regressions: Innovation Quantity | 23 |
| 6 | Cross-CPC Regressions: Innovation Quality | 24 |
| 7 | Early Ending Samples | 26 |
| A1 | First Digit CPC Code | 33 |
| A2 | Example of CPC subclasses from one patent | 33 |
| A3 | Example of Country-CPC weight methodology for one patent | 34 |

# 1 Introduction

Given the seminal work of endogenous growth theory by Romer (1990), it is widely accepted amongst Economists that innovation is a fundamental driver for economic growth. But innovation requires human and physical capital to conduct stimulating research and inventions (Aghion & Howitt, 1992). Therefore, innovation seldom occurs independently. This has led to the introduction of direct fiscal incentives such as R&D tax credits to increase firms' tax reliefs (Hall & Van Reenan, 2003; Bloom, Griffith & Van Reenan, 2013). As a result, in theory, firms will use its tax relief surplus for innovation and research through investment in human and physical capital.

Though direct fiscal incentives on innovation is of great significance for stimulating innovation and economic growth, in this thesis, we contribute to the academic literature by exploring the indirect fiscal incentive of corporate tax and their effect on innovation activity and the quality of innovation.

A negative relationship between corporate tax and innovation tends to be the most dominant finding in the literature: as corporate tax rises, firms' after-tax income falls, yielding less room to invest in innovation. However, previous literature on corporate tax and innovation that focuses on European data has conducted an extensive amount of research toward their contemporaneous relationship at a firm-level. From this, we propose the following questions: 1) are the effects truly contemporaneous? and 2) does the inverse relationship hold when we examine it at a cross-country setting?

To address the first question, we can theorize that the effects of corporate tax on innovation may not occur contemporaneously: new inventions are driven by failure and learning-by-doing along with some serendipity. Thus we may not witness the outcome of innovation output in the same year. To take account of the inertia of innovation, we follow Akcigit, Grigsby, Nicholas and Stantcheva (2022) amongst others by replacing the contemporaneous corporate tax with a three year lag.

However, our initial dataset contain missing variables that help examine the relationship between corporate tax and innovation at a firm-level. We source patent data from the European Patent Office's (EPO) database called PATSTAT. This original dataset consists of a cross-section of patent applications from 1978 to 2019. This dataset does not provide information on firm-level controls such as firm size and firm age. Due to this limitation, this thesis will examine innovation's inertial behaviour at a cross-country setting.

Adopting a cross-country analysis in this thesis allow us to compare our results with Bosenberg and Egger (2017) and Ernst, Richter and Riedel (2013) who examine the contemporaneous relationship between corporate tax and innovation output, also at a cross-country setting. To measure innovation output, we use patent quantity and patent quality as our dependent variables to proxy for a country's innovation activity and innovation quality, respectively.

Using a Pseudo Poisson Maximum Likelihood model, we found from our baseline results that given a 10% increase in corporate tax three years ago, on average, we can expect a country's innovation activity and quality of innovation to fall this year by 0.83% and 0.64%, respectively. These coefficients are statistically significant at the 5% level for both innovation quantity and innovation quality. In addition to our baseline results, we also find that an increase in corporate tax this year will negatively effect a country's innovation quality up to the third year and will have no statistically significant effect from the fourth year on. Innovation activity on the other hand, is not effected until the third year on. We speculate that a country's innovation activity and quality may be dominated by multinationals. With greater financial flexibility, multinationals can relocate their patents in an alternative country more easily than domestic firms given an increase in the host country's corporate tax rate (Almodovar & Nguyen, 2022).

Along with these results, we implemented alternative aggregations to examine the heterogeneity in innovation across the quality of inventors and technology classses. We found that the alternative micro-aggregation of solely inventors with patent citations falling in the top 5%, or "star" inventors as defined by Moretti and Wilson (2017), are more elastic to changes in corporate tax than the micro-aggregation of only non-star inventors; and the statistically significant inverse relationship between corporate tax and innovation is consistent across most technology classes.

This thesis contributes to two strands of the literature. The first is taxation and economic growth. There is a vast amount of research regarding tax policies on economic growth (Romer & Romer, 2010); firm entry (Gentry & Hubbard, 2005); entrepreneurial risk taking (Djankov, Ganser, McLiesh, Ramalho & Shleifer, 2008); and location decisions (Moretti & Wilson, 2017). Similar to this thesis, Bosenberg and Egger (2017) uses 106 countries and found a negative contemporaneous effect between corporate tax and innovation activity at a cross-country setting. The authors however, ignore the inertial behavioural of innovation. On the contrary, with US state-level data for corporate tax and micro-aggregates for innovation levels, Akcigit et al. (2022) found a negative relationship for a three year lag in corporate tax and innovation activity and innovation quality across US states. In this thesis, we adopt the econometric specification by Akcigit et al. (2022) and apply this to a cross-country setting for countries who are members of the European Union. We use country members of the European Union to follow Ernst et al. (2013) as closely as possible, who also found a negative contemporaneous relationship between corporate tax and innovation.

The second strand of literature we contribute to is the determinants of innovation. Previous literature has shown that firm-level innovation is heavily determined by competition (Aghion & Howitt, 1992); financial development and regulation (Hsu, Tian, & Xu, 2014); managerial incentives (Manso, 2011); and attitude toward failure (Tian & Wang, 2011). We contribute to this strand of literature by showing that corporate tax is also an important determinant of innovation, even if we are examining this across countries.

The remaining section of this thesis is as follows. Section 2 reviews the literature on the measurements of innovation and corporate tax and innovation. Section 3 explains how we cleaned and derived our patent data along with the gathering of cross-country variables. Section 4 presents the baseline results and robustness checks. Section 5 concludes.

# 2 Literature Review

## 2.1 Innovation Measurements

The two most common measures of innovation in the literature are patents and R&D expenses.

R&D expenses measures innovation inputs rather than innovation outputs. Therefore, high R&D expenses may not necessarily derive innovation of high quality. For example, assuming a dynamic game of complete information, Krieger (2021) found that the propensity to terminate R&D projects increases when firms in the biopharmaceutical industry receive news that a similar project of their competitor has failed. For the firm who received the news of failure, this yields high innovation input but low innovation output. Another example on input not equating output is Nokia and Apple during the fiscal years from 2001 to 2011. Nokia had three times more R&D expenses than Apple, yet still lost its competitive position from introducing irrelevant market products (Lerner & Seru, 2022). Moreover, R&D expenses does not measure the input for one invention or geographic location: rather, it explains innovation effort at a firm-wide level. This restricts researchers in understanding the heterogeneity in innovation at a deeper level, casting further doubt for using R&D expenses to analyse innovation at a cross-country setting.

As mentioned in the introduction, innovation is an important driver for growth, and hence the introduction of R&D tax credits: as firms attain a certain threshold of R&D expenses, they pay less tax, increasing after-tax profit and thus incentivizing firms to recycle this surplus into innovation. However, with the presence of information asymmetry, Chen, Liu, Serrato and Xu (2021) found that 24% of Chinese firms agglomerate miscellaneous expenses into their R&D expenses to achieve the tax credit threshold. Firms may also use the surplus from R&D tax credit to increase salaries and wages rather than innovation activities (Goolsbee, 1998). Thus, given the amount of drawbacks for R&D expenses, we will not pursue this measurement for innovation.

Therefore, this thesis uses patents as an alternative measure for innovation. A patent is a licence for the inventor to contain ownership of a good or service. An inventor files for a patent with the incentive of gaining a monopoly on their invention for a certain time. This prevents competitors or other inventors from stealing their work during this period. Patents are a good measure for innovation because it provides information on the quantity and quality of innovation, or the output of innovation. Although a large quantity of patents filed help explain innovation activity, it may not necessarily imply these patents are of high quality. Innovation quality is proxied by the number of cites a patent receives (Trajtenberg, 1990; Hall, Jaffe & Trajtenberg, 2001). The general hypothesis is that higher levels of citation implies other inventors are interested in this patent, subsequently building upon it.

Patents of high quality tend to be dominated by a small group of inventors, or the top 5%, also known as star inventors (Moretti & Wilson, 2017). These inventors usually work for multinationals, as multinationals have the finance to fund these highly skilled and valuable economic agents. Analysing the levels of innovation between domestic and multinational firms in Spain from 2006 to 2010, Almodovar and Nguyen (2022) found that domestic firms innovate less than their foreign subsidiaries. As patents contain information on the location of an inventor, the academic literature has found that multinationals locate where high skilled labour are abundant (Akcigit, Baslandze & Stantcheva, 2016; Moretti & Wilson, 2017).

Despite its ability to tackle the drawbacks of R&D expenses, patents as a proxy for innovation are not a means to an end. One problem is truncation bias (Hall et al., 2001): a patent filed at the beginning of the sample will have more cites than a patent filed toward the end of the sample as the latter has less time to accumulate more cites. For example, given a sample that runs from 2010 to 2020, a patent filed in 2010 may have more cites than a patent filed in 2020 and thus the ability to measure the quality of the latter invention diminishes. From this result, the sample suggests higher quality innovations are clustered around earlier periods, which could be erroneous.

Another component of patents to consider is the heterogeneity in patenting activity across technology classes: the rate of patenting across technology classes differ. Acemoglu (2002) attributes the difference in technical change to the market size effect and the price effect.

The market size effect explains that the size of the market demand is positively correlated with the direction of innovation for this market. Acemoglu and Linn (2004) examined the relationship between patent intensity in the pharmaceutical industry and the market size, measured by consumer demographics. They attribute a positive relationship between current and future market size and innovation. Firms in pharmaceutical research can predict the future consumer segment required for medication through data on current demographics and subsequently adjust their innovation based on this variable. Costinot, Donaldson, Kyle and Williams (2019) provide a similar result at a cross-country setting whilst taking account of the home market effect (where countries specialise in and export more of a certain good if domestic demand is high). The authors found that countries invest more and export more of specific drugs when domestic demand for these drugs are high.

Aghion, Dechezleprêtre, Hemous, Martin and Van Reenan (2016) presents a similar finding but within the strand of environmental economics. They examine the effects of clean and dirty innovation for the automobile industry given an increase in carbon taxes for firms and consumers. An increase in carbon tax reduces consumer demand for dirty cars, or cars with internal combustion engines. This therefore leans consumers toward a higher demand for electric and hybrid cars. The authors found that an increase in carbon tax directs innovation toward clean automobile patents, and away from dirty automobile inventions.

The price effect on the contrary, explains the elasticity of substitution for factor inputs: if price of factor input rises, inventors will direct their innovation toward more efficient inventions.

During the British Industrial Revolution, Britain witnessed a cotton shortage due to the advent of the US civil war, leading to an increase in the price of cotton. The technologies for cotton in Britain were adapted to US cotton, as US cotton were of lower cost and higher quality relative to other cotton varieties such as India, Egypt and Brazil.[1] Hanlon (2015) finds that the

[1] Hanlon (2015) explains that non-US cotton tend to have passed through the hands of multiple intermediaries, who added salt and water to increase the weight of cotton. As a consequence, these cottons are more dirty and requires more effort to clean.

blockade of Southern shipping in the US has directed patenting in textile toward the non-US cotton, whilst patenting for non-textile sectors remained constant.

Moscona and Sastry (2023) examines the technological response in agriculture in the US for climate change since the mid 20th century. The authors found that agricultural innovation and patenting has been directed toward crops that are more exposed to extreme climate conditions. As a result, the effects of climate change on these crops have been mitigated.

Therefore, patent intensity across technology classes is driven by both market size and the price of factor inputs. Nevertheless, patent-intense technology classes will thus have higher count of patent forward citations, but this could be inflated due to a larger quantity of patents applied. This further implies a heterogeneity in the rate of patent obsolesce across technology classes, or the propensity to cite older patents. Lastly, inventors usually cite prior patents that help contribute to their invention. However, this inventor may cite their own work. This inflates the number of citations for their previous patent which artificially increases the quality of the invention (Hall et al., 2001).

Despite these disadvantages for patents, their benefits tend to outweigh R&D expenses. Furthermore, this thesis is interested in examining the quantity and quality of innovation, an element R&D expenses fails to capture.

## 2.2 Corporate Tax and Innovation

What are the contemporaneous effects of corporate tax on innovation? Bosenberg and Egger (2017) examined the cross-country relationship between corporate tax and innovation. Their findings align with the negative relationship between corporate tax and innovation. Using yearly micro-aggregate data for granted patents from 1996 to 2012 for 106 countries, they found that an increase in the effective average tax rate (EATR) contemporaneously reduces patent activity.

Cai, Chen and Wang (2018) observed similar results with yearly Chinese firm-level data from 1998 to 2007. They implement a Regression Discontinuity Design for the introduction of a corporate tax reform in China in 2002: firms established after this reform pays more tax. The aim of this policy is to stimulate competition and encourage firms entered post-2002 to innovate. They found the introduction of this policy increased patent quantity and quality as firms rely on after-tax profit for innovation activity. He, Jiang and Frand (2023) apply the same methodology as Cai, Chen and Wang (2018) but on a different corporate tax reform in China in 2012 and found similar results. He, Jiang and Frand (2023) found a negative effect between corporate tax and innovation but the effect was greater for small firms, financially constrained firms and high-tech firms.

With European micro-aggregate data of granted patents from 1995 to 2017, Ernst, Richeter and Riedel (2013) found that corporate tax has an inverse relationship with innovation quality. They observed that a 10% reduction in the effective corporate tax rate increases patent quality by 5.6% in the same year. The authors attribute the negative effect to multinational corporations shifting innovation output to low tax countries to attain a higher after-tax profit. As a consequence, countries with lower corporate tax rates yield higher quality innovation. Griffith, Miller and O'Connell (2013) and Karkinsky and Riedel (2012) found similar results to Ernst, Richeter and Riedel (2013). Griffith et al. (2013) uses data on firms that have applied for a patent at the EPO from 1985 to 2005 and observed that corporate tax increases can lower patent activity but increase government revenues. Karkinsky and Reidel (2012) used multinational corporations of 18 country members of the European Union from 1995 to 2003 and found that patenting activity is negatively related to the host country's corporate tax rate. Along with this, both Griffith et al. (2013) and Karkinsky and Rieldel (2012) attribute the fall in patenting activity to the profit shifting of multinational corporations.

This attribution can be supported by Cheung, Guo, Weng and Wu (2021) who use US firm level data from 1987 to 2012 and found that corporate tax planning and innovation are related. More specifically, corporate tax planning for small domestic firms have no effect on patenting, but has an impact on multinationals as they have the flexibility to shift their intangible assets to low-tax countries. Findings by Cheung, Guo, Weng and Wu (2021) align with Grubert and Slemrod (1998). With data on tax return files of 214 firms in 1987, Grubert and Slemrod (1998) found that Peurto Rico's highest patent industry was pharmaceutical and electronics, which came from multinationals located in Peurto Rico for corporate tax benefits.[2]

Howell (2016) however, found different results on the relationship between corporate tax and innovation. Howell (2016) used Chinese firm-level data from 2001 to 2007 and found that the effects on changes in corporate tax were greatest amongst financially constrained firms, where they proxy financial constraint with the company's cash flow. Although Howell (2016) found that a change in corporate tax has a significant effect on the sales of company, it had no impact on innovation. A zero effect of corporate tax on innovation align with the findings of Chen and Miller (2007). Chen and Miller used publicly traded US data from 1980 to 2001 and observed that tax relief went toward sales and marketing rather than innovation.

Shakhmuradyan (2022) had similar findings with eleven countries that are located in Eastern Europe from 2010 to 2019. Rather than using patents as their dependent variable and the effective tax rate as their variable of interest, they measure innovation with Business Expenditure on Research and Development and corporate tax as Total Tax and Contribution Rate. They found no effect of corporate tax on innovation, but rather innovation for any given country in their sample is mostly explained by the number of researchers in R&D, trade openness and college education. This statistical insignificance of corporate tax may be attributed to using an alternative measurement for innovation. Another possibility is that the innovation activity in Eastern Europe is relatively low compared to other regions of Europe (European Innovation Scoreboard, 2023). As a result, low levels of innovation output imply low innovation input and therefore a lack of variation in innovation input. Thus corporate tax does not covary with innovation.

But are the effects of corporate tax on innovation truly ascribed contemporaneously? Inno-

[2] Thus, given that low corporate tax rates attract multinationals who produce patents of higher quality, some countries in Europe have introduced alternative corporate tax policies to attract patents such as patent boxes (see Alstadsæter, Barrios, Nicodeme, Skonieczna and Vezzani (2018) for an analysis on patent box and their effect on patent quantity and quality; and see Davies, Kogler and Hynes (2020) on the probability of a patent being granted under the presence of a patent box) and the B-Index (see Warda, 2002).

vation and patenting is a learning-by-doing process, implying that inventions may not take place in the same year. Therefore, we may not witness the effect of corporate tax on innovation until a few years later.

Akcigit, Grigsby, Nicholas and Stantcheva (2022) take this into account by analysing the relationship between EATR and innovation at a state-level in the US. With yearly state-level tax data and micro-aggregate for innovation data from 1940 to 1999 in the US, Akcigit et al. (2022) estimated that a 1% fall in EATR significantly increased innovation quantity, innovation quality, the number of inventors and share of inventors employed by companies by an average of 2.8%, 2.4%, 2.3% and 0.6% after three years, respectively.

To take account of the heterogeneity in the effects of EATR on innovation across firms in US states, Mukherjee, Singh and Žaldokas (2017) assigned dummy variables for increases and decreases in EATR. They gather data on 47,632 patents from firms in the US from 1990 to 2006 and found that a decrease in the EATR had no effect on innovation, whilst an increase did. Atanassov and Liu (2020) found similar results to Mukherjee et al. (2017) with the same econometric methodology. Collecting annual data from 1988 - 2006 of 8,013 US firms, they found that both an increase *and* decrease in the effective corporate taxes had an effect on innovation.

The possible reason for the difference in findings between Mukherjee et al. (2017) and Atanassov and Liu (2020) is that the former used a dummy for all increases and decreases in corporate taxation. From this, there is a lack of variation in corporate taxes due to a large amount of dummies. Atanassov and Liu (2020) on the other hand, introduced a dummy when there was a large corporate tax increase or decrease, where they define large as 1% or more.

The literature on corporate tax and innovation leans heavily towards an inverse relationship, irrespective of accounting for innovation's inertial behaviours. However, literature on the inertial effects of corporate tax on innovation are mainly attributed to the US: we test whether this inertial relationship holds at a cross-country setting for members of the European Union.

# 3 Data

Dissected into patent data and country data, we will explain why we selected certain variables in this section. Furthermore, we delineate each step on how we cleaned and obtained our final micro-aggregate dataset.

## 3.1 Patent Data

The patent data was sourced from the EPO database called PATSTAT. The original dataset is based on a cross-section of each individual patent application from 1978 to 2019. Each application contain the following variables: patent application ID; the patent's earliest filing year; the patent's granted year; a binary yes/no status on whether the patent was granted; the inventor ID for each patent; the country each inventors are resided in; the country residence of the patent applicant; each patent's technology class; the count of forward citations; and the count of backward citations. This initial dataset had 3,739,897 patent applications.

As the purpose of this thesis is to examine innovation at a cross-country setting with EU members from 2000 to 2017, we first select the appropriate time period. The time series running from 2000 to 2017 deems the most feasible years in terms of collecting data from alternative sources. 2000 and 2017 were the minimum and maximum year where data was available for all variables, respectively. Note the years we have selected is based on the earliest filing year of the patent application: firms are eager to protect their inventions as soon as possible and therefore it is the closest year to the actual creation of the invention. Publication year of the patent application can deviate across each application and technology class as this decision depends on the EPO and is therefore not helpful in explaining initial patent inventions.[3]

To account for the cross-country setting, we use the inventor's country of residence. We avoid the patent applicant's country of residence because we can have multiple inventors contributing to the one invention in which they may reside in different countries. After selecting patents with inventors located in the EU, and a time series from 2000 to 2017, we now have 946,825 patent applications.[4] Malta contains missing tax variables (see section below) and we therefore omit patent applications applying from Malta, or 164 applications. This brings us to 946,639 applications. Not all 946,639 applications were granted a patent. In fact, 453,524, or 48%, of these applications were rejected from a patent. Nevertheless, we keep these applications in our dataset as it is still possible to cite non-granted patent applications. Although there tends to be a difference in quality between granted and non-granted patents, inventors do not know whether their patents will be granted or not. This will help further explain a country's propensity to patent under the presence of uncertainty. Patents may be rejected for their lack of novelty in their invention. However, solely using granted patents will give more weight to a country's innovation quality than it should. Therefore, in this thesis, we examine patent quantity and patent citations of both granted and non-granted patent applications.

As mentioned, it is possible to have multiple inventors for the one patent, and inventors for this one patent may be residing in multiple countries. Therefore, in order to avoid double counting innovation levels, we fractionally weigh the invention conditional on the inventors residence location. For example, if there were two French inventors and three Spanish inventors for the one patent, France will have a fractional weight of 0.4 and Spain a fractional weight of 0.6, where the weights sum to 1.[5] We define this as a country-weight.

However, by solely weighing our innovation levels conditional on countries ignores valuable information on the heterogeneity of the effects of corporate tax on innovation across technology classes. To take account for this heterogeneity, as patents are not classified by industry but by technology classes, we use the EPO's Cooperative Patent Classification (CPC) system and fractionally weigh each patent with respect to CPC.[6] CPC in our dataset is a four digit variable to explain which class the patent falls under. Furthermore, a patent may have multiple CPC's to represent the brevity of an invention. Similar to countries, we fractionally weigh the number of

[3]It takes on average 5.5 years for a patent to be granted.

[4]We did not use residents of the United Kingdom due to Brexit.

[5]Unless some of the inventors for a patent are not resided within the EU, then the sum is less than one.

[6]A table of the 9 first-digit CPC codes are provided in the Appendix of this thesis.

distinct CPC's a patent has, or a CPC-weight.

So far, we have two independent weights: a country weight and a CPC weight. We want to include the heterogeneity of a country's innovation output across technology classes. We extend on the country weight by multiplying it with the CPC weight. This becomes a Country-CPC weight.

In short, aggregating across inventors for a given country with respect to Country-CPC weight or solely country weight derives the same innovation statistics. Country-CPC weight only extends country weight in that it accounts for technology classes which allow us to examine the heterogeneity across each country's patent specialization. Solely using country weight fails to capture this element because we are ignoring the technology class variable. We provide an explicit example on the methodology for a cross-country-CPC weight for one invention in the appendix.

Although patent quantity tends to be positively correlated with the number of forward cites, patent quantity does not conclude higher levels of innovation (Hall et al., 2001). It is quite common to have a low number of forward citations. In fact, prior micro-aggregating our dataset, 668,774 (71%) patent applications have zero forward citations.[7] 283,842 of these patent forward applications are from less than the median year of 2009, or 43% of the total 668,774 patents with zero forward citations. This suggests that the remaining 57% of patents with zero forward cites *may* be attributed to the truncation bias problem. As discussed, even though forward citation count is a good measure of innovation quality, a common problem is the sample truncation. As a reminder, this is when a patent filed earlier in the sample contains more cites than one filed towards the end of the sample as the latter has less time to accumulate cites.

To illustrate this, we dissect the dataset into two alternative aggregations which follows the Country-CPC weight methodology. The first subset micro-aggregates solely inventors whose patents contain cites that fall in the top 5%, or the star inventors. The second subset is a micro-aggregate of inventors whose patent citations belong to the bottom 95% of citations. The difference in the star and non-star inventors demonstrate the nonlinearity in innovation. Figure 1 illustrates truncation bias for the two alternative micro-aggregates.

[7]As a side, backward citations counts the number of previously cited patents for a patent. This helps explain the importance of the prior art and subsequently how inventors diffuses knowledge.

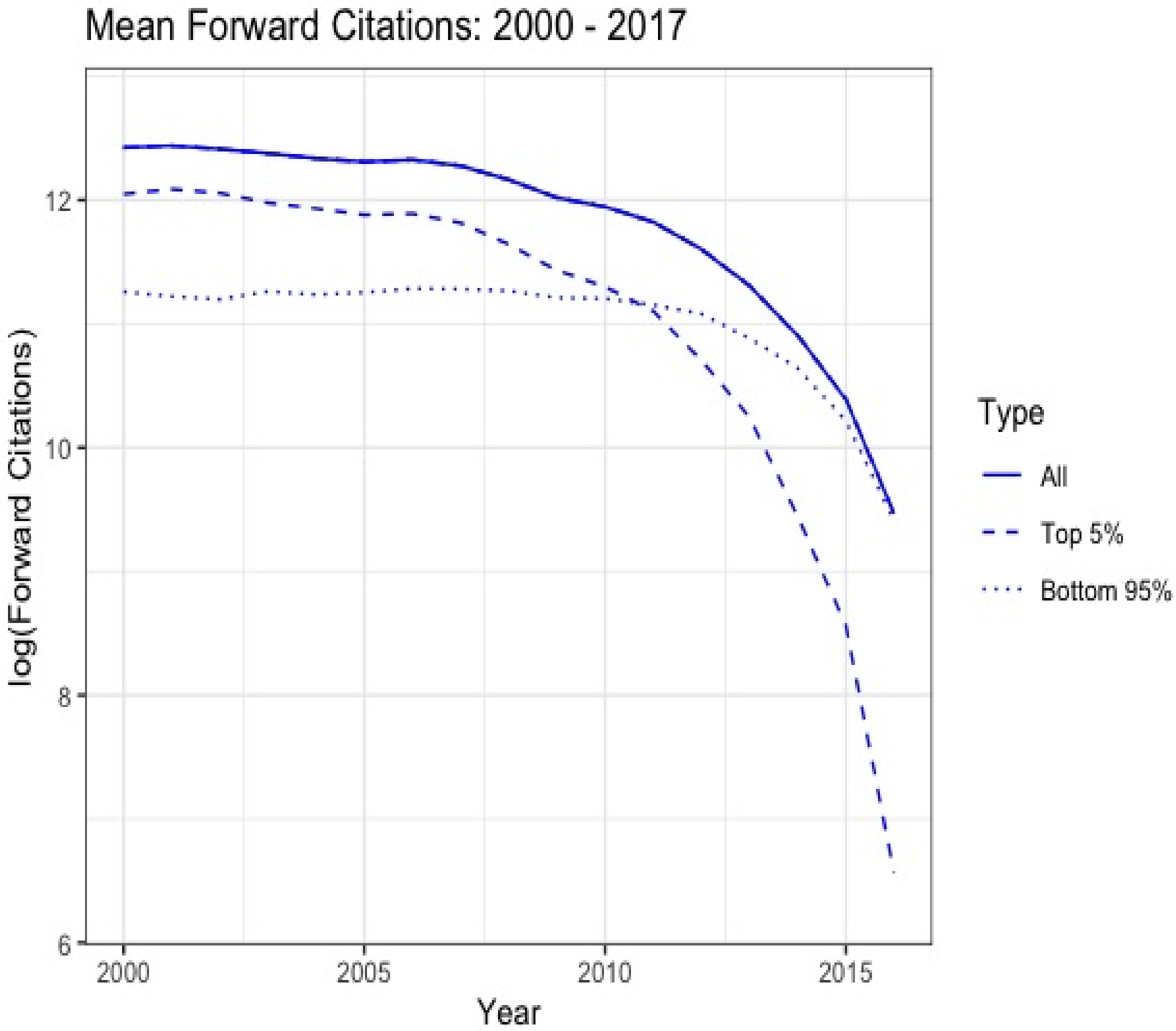


Figure 1: Forward Citations and Truncation Bias

From Figure 1, micro-aggregates of solely star inventors are closer to the total citations than micro-aggregates of only non-star inventors in the earlier periods. However, the divergence between the star-inventor citations and the total citations occur around 2013.

The initial relationship explains that provided patents have time to accumulate cites, the aggregate quality of patents in a country are heavily dominated by a small group of inventors, whilst the remaining inventors have a marginal contribution. Although we do not have data for this, we could theorise that these remaining inventors work for domestic firms, firms who, according to Almodovar and Nguyen (2022), tend to patent less and also have less forward citations.

Toward the end of the sample, the quality in innovation for a country who produces patents less than the top 5% is greater than countries with the aggregation of solely star inventors. This can be attributed to the fact that the quantity of patents with cites below the top 5% of citations is greater than patents with cites above the top 5%. Therefore the non-star inventors seemingly yield higher innovation quality. As truncation bias could effect our results, we will implement a robustness check for this specific problem.

## 3.2 Country Data

We follow Akcigit et al. (2022) and others by using the EATR. The EATR data was calculated by Kostarakos and Varthalitis (2020) for country members of the EU.

Using data from the Eurostata database, they calculate EATR as the ratio of tax revenue to tax base. For corporate EATR, they measure this with:

$$\tau_{it}^{corp} = \frac{tax\ revenues_{it}}{tax\ base_{it}}$$

Where the numerator is the sum of total taxes paid by firms for a given country $i$ at year $t$, leading to government tax revenues. The denominator is the sum of every company's pre-tax earnings. This is known as the tax base.

We control for other cross-country predictors that help explain the deviations in innovation, all sourced from the World Bank Dataset.

We follow Akcigit et al. (2022) by controlling for GDP Per Capita, R&D to GDP ratio, and population growth. GDP Per Capita, measured in US dollars, explains the wealth of nations and the institutional environment. If GDP Per Capita rises, we can expect innovation output to increase also. We include R&D to GDP ratio to control for the R&D intensity of a country. A high innovation input should, in theory, yield high innovation output. Lastly, to control for the size of the economy, we include population growth which, along with country fixed effects that absorb the initial value of population, controls for the number of people. Population of a younger cohort tend to be more innovative relative to older cohorts (Packalen & Bhattacharya, 2019). Moreover, a younger population implies greater tax revenue in which governments will finance toward the country's R&D intensity. However, an increase in innovation occurs when population growth is around 0.75%, else innovation falls (Coccia, 2014).

Following Shakhmuradyan (2022), we also control for the number of R&D researchers, the openness of a country and college education (tertiary education) in a country.

Openness of a country, proxied by the sum of Import to GDP ratio and Export to GDP ratio (becoming Trade to GDP ratio), argues that as a country becomes more open, we should expect innovation to increase. According to Grossman and Helpman (1990), when a country becomes more exposed to foreign trade, spillovers in foreign technological knowledge from advanced countries will occur. Domestic workers absorb and diffuse this knowledge which consequently yield higher levels of innovation through learning-by-exporting.

However, the diffusion and accumulation of knowledge depends upon human capital (Aghion & Howitt, 1992). Thus, as per Wang (2010), we include the number of researchers in R&D to control for the scale of the R&D sector; and control for human capital stock with college education. Through the lenses of creative destruction, these variables diffuse knowledge more rapidly which reduces monopoly rents by turning innovations obsolete until the next innovation occurs.

Lastly, we control for the observed and unobserved variation across time through the inclusion of year dummies. These dummies filter out the effects of common shocks.

### 3.3 Summary Statistics

Table 1 below presents the summary statistics in levels-form for our innovation output variables, and in natural logarithmic-form for our cross-country controls. This table is based on the final dataset at country level where we have a full balanced panel.

Table 1: Summary Statistics

| Statistic | N | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|
| Year | 468 | 2,008.50 | 5.19 | 2,000 | 2,017 |
| Quantity (Patent) | 468 | 1,928.74 | 4,558.33 | 2.66 | 24,614.71 |
| Quality (Patent) | 468 | 2,664.36 | 7,804.32 | 0.00 | 59,204.98 |
| EATR | 468 | −1.72 | 0.52 | −3.78 | −0.16 |
| GDPPerCap | 468 | 9.98 | 0.81 | 7.39 | 11.73 |
| PopulationGrowth | 468 | −0.59 | 0.95 | −4.62 | 0.89 |
| R&D-GDP | 468 | −4.44 | 0.65 | −6.09 | −3.25 |
| Openness | 468 | 0.003 | 0.43 | −0.79 | 1.26 |
| Researchers | 468 | 7.87 | 0.55 | 6.07 | 8.97 |
| College | 468 | 4.07 | 0.39 | 2.18 | 4.90 |

Quantity and Quality of patents are the micro-aggregates of each patent's Country-CPC quantity and forward citations. For any given country in our dataset, the mean of GDP Per Capita in US dollars in logarithmic form is 9.98. The logarithm of population growth, or the log of the percentage change in population relative to the previous year, is on average -0.59, but also witnesses higher variances. All observations are negative for R&D to GDP in logarithmic form. This is because R&D to GDP ratio is less than unity. Though openness of a country tend to be quite close to zero, it also has a great range. This is attributed to Luxembourg with a high level of openness. Prior taking logs, the number of researchers in R&D was initially measured by per million people and College was the ratio of people in tertiary education relative to the gross number of people in that age cohort. The EU has 27 countries. But as aforementioned, we exclude Malta due to missing tax observations. We therefore have 26 countries for a time series of 18 years, or 486 observations for our final micro-aggregate panel data.

In any given year, the average number of patent quantity for a country in our dataset is 1928.74 and the average number of aggregate forward citations, or quality, is 2664.36. Furthermore, these two variables tend to deviate from the mean quite drastically, as shown by the standard deviation. The significant amount of variation help elucidate whether innovation co-vary with other controls.

Some years however, a country's forward citation could be zero. To be exact, ten observations in our dataset have zero citations. There are two possible reasons for this. The first is the truncation bias as previously mentioned: a patent may have been filed at a period toward the end of the sample and thus has less time to accumulate cites. Consequently, aggregating patents at the end of the sample to a cross-country setting does not change this result. In fact, nine of the ten observations with zero cites were in the years 2016 and 2017, justifying the end of sample truncation.

Another reason is that the patent was filed early in the sample, but had zero cites because it did not intrigue other inventors. As discussed, it is not uncommon to find a patent with zero cites as other inventors may not find them of interest. Cyrus in 2003, the remaining observation without any forward citations, fell victim to this.

The natural logarithm of EATR varies quite significantly: we have a standard deviation of 0.52 and a range of 3.94. The high spread and range may be attributed to the importance of indirect fiscal incentives for innovation: as time progresses, EATR falls in order to indirectly incentivize firms to conduct further research in innovation. Our plot in Figure 2 with the natural log of EATR from solely 2000 and 2017 supports this hypothesis, where we see the density for EATR is greater on average in 2000 than 2017.

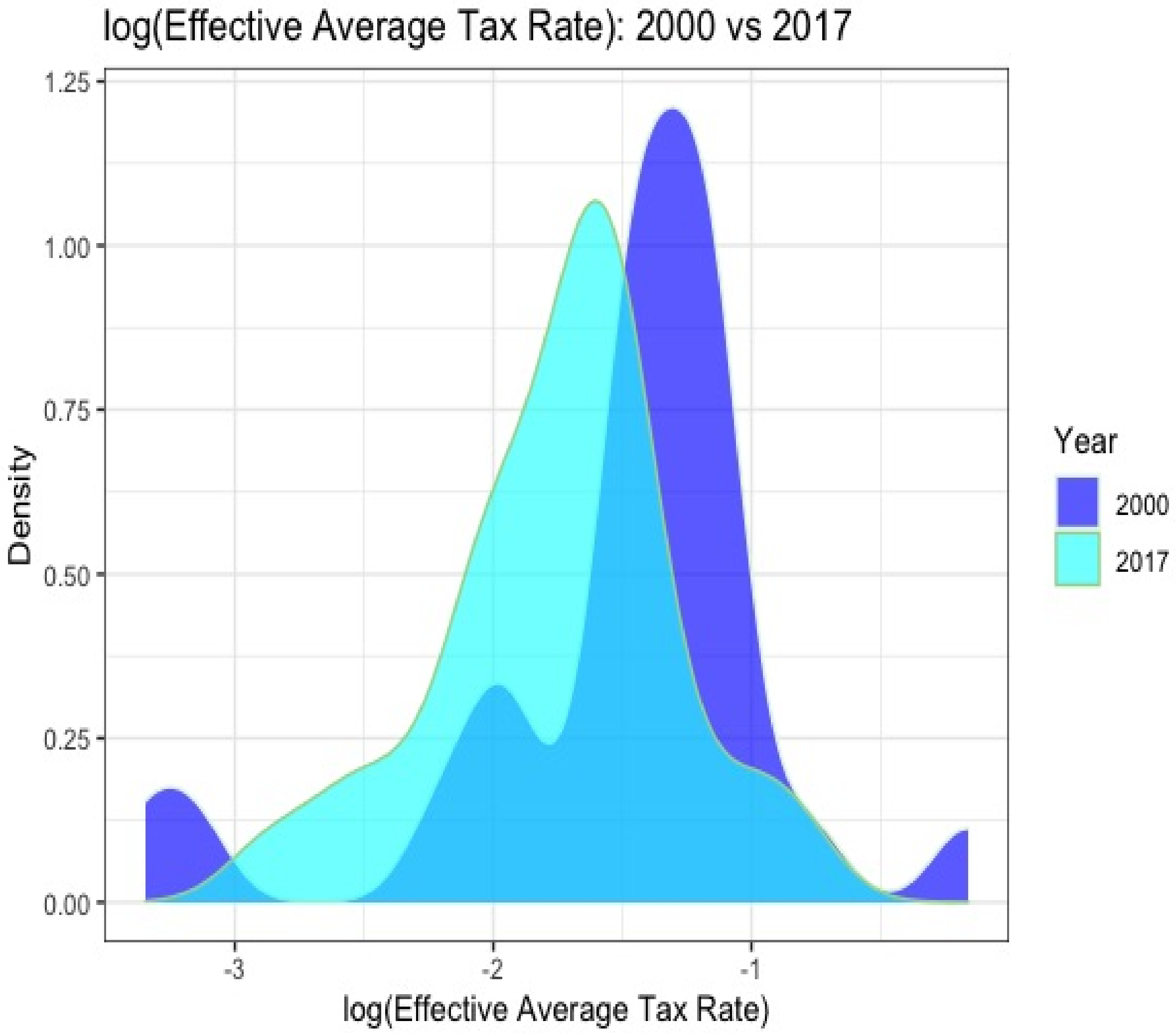


Figure 2: Effective Average Tax Rate in 2000 and 2017

### 3.4 Econometric Methodology

As our dependent variable is a count variable, we implement a Psuedo Poisson Maximum Likelihood (PPML) regression. Unlike an OLS fixed effect, PPML takes account the possibility of zeros in our dependent variable, and consequently does not omit this observation when we transform it into logarithmic form. The PPML is specified as:

$$Y_{zit} = exp(\beta_{EATR} ln(EATR_{i,t-3}) + \mathbf{X}'_{it}\gamma + ln(\delta_i) + ln(\delta_t))\varepsilon_{it} \tag{1}$$

Where $z \in \{Quantity, Quality\}$. $\beta_{EATR}$ is my coefficient for EATR. This coefficient for my variable of interest explains the elasticity of corporate tax and innovation. The null hypothesis at the 5% level states that corporate tax has an effect no different than zero on innovation, whilst the alternative says it has an effect different to zero. Thus, this coefficient examines a 1% increase in EATR three years ago and the percentage outcome on innovation output this year.[8]

We select this lag length to follow Akcigit et al. (2022), Mukherjee et al. (2017), and Atanassov and Liu (2020) in order to account for innovation's inertial behaviours. Though this number is somewhat arbitrary, we test the robustness of other lag lengths in section 4.2.1 in this thesis. The introduction of a lag on our variable of interest avoids the endogeneity problem: a shock in our error term this year cannot influence previous corporate tax rates.

$\mathbf{X}'_{it}$ is a vector of cross-country controls in logarithmic form, namely GDP Per Capita, R&D to GDP ratio, population growth, number of researchers, openness of a country and college education.

To control for the variation across time and space, we include fixed effects dummies $\delta_i$ and $\delta_t$ for country and year, respectively. We also cluster the standard errors at the country level: countries who have the same economic characteristics tend to have similar but differentiated innovation behaviours (Linder, 1961). Therefore each country's innovation output may not be independent of one another. Lastly, we include an exogenous shock term $\varepsilon_{it}$ to take account of the unexplained and unobservable variables.

[8]Note a 1% increase in $EATR_{t-3}$ is not $EATR_{t-3} + 1\%$, but rather $EATR_{t-3} \times 1.01$.

# 4 Results

## 4.1 Baseline Results

In the table below we present our baseline results with the PPML regression.

Table 2: Baseline PPML Regression

| | Quantity$_t$ | | | Quality$_t$ | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| EATR$_t$ | -0.018 | | | -0.033** | | |
| | (0.078) | | | (0.015) | | |
| EATR$_{t-3}$ | | -0.125** | -0.083** | | -0.089** | -0.064** |
| | | (0.064) | (0.041) | | (0.042) | (0.031) |
| GDPPerCap$_t$ | 1.018*** | | 0.846*** | 1.773*** | | 1.660*** |
| | (0.311) | | (0.289) | (0.402) | | (0.313) |
| R&D-GDP$_t$ | 0.569*** | | 0.480** | 0.521** | | 0.768*** |
| | (0.205) | | (0.233) | (0.223) | | (0.249) |
| PopulationGrowth$_t$ | -0.152*** | | -0.164*** | -0.124 | | -0.141** |
| | (0.057) | | (0.055) | (0.078) | | (0.056) |
| Openness$_t$ | -0.584*** | | -0.450* | -0.124 | | -0.399 |
| | (0.152) | | (0.236) | (0.207) | | (0.310) |
| Researchers$_t$ | -0.263** | | -0.278*** | -0.277 | | -0.503*** |
| | (0.103) | | (0.098) | (0.182) | | (0.181) |
| College$_t$ | 0.286** | | 0.194 | 0.292* | | 0.201 |
| | (0.129) | | (0.142) | (0.157) | | (0.192) |
| Observations | 468 | 390 | 390 | 468 | 390 | 390 |

Robust standard errors in parentheses. All specifications include country and year fixed effects.
Robust standard errors clustered at country level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To first test whether the contemporaneous effects of EATR on innovation for our cross-country setting has a similar behaviour to country-level results from the literature, columns 1 and 4 estimates the contemporaneous effects of EATR on innovation whilst controlling for other cross-country variables. We see that a 10% increase in EATR this year reduces innovation quantity by an average of about 0.20% in the same year, holding all else constant. However, EATR here is statistically insignificant. We therefore fail to reject the null that EATR has an effect no different to zero on innovation quantity. Our results contradict Bosenberg and Egger (2017) who find a negative contemporaneous cross-country relationship between EATR and patent quantity with micro-aggregates on granted patents. The possible explanation for this difference may be that Bosenberg and Egger (2017) used a much larger cross-section of 106 countries, relative to ours of 26. We may also speculate that firms do not have time to respond to an increase in EATR in the same year, and hence there is no statistically significant effect in innovation activity.

Innovation quality on the contrary, has a slightly greater effect in absolute terms, as a 10%

increase in EATR reduces a country's innovation quality in the same year by an average of 0.33%. This result is statistically significant at the 5% level. This effect in absolute terms is smaller than the effect size of 5.6% by Ernst et al. (2013). The reason for this discrepancy may be because Ernst et al. (2013) restricted their micro-aggregate dataset to granted patents, whilst our micro-aggregate consists of both granted and non-granted patents. As discussed, there tends to be a significant difference in quality with respect to granted and non-granted patents, where non-granted patents tend to contain much lower cites. Consequently, by ignoring non-granted patents, Ernst et al. (2013) found results greater in absolute value.

In summary, our contemporaneous results here align with the negative relationship between corporate tax and innovation quality from the literature, but differs with regards to corporate tax and innovation quantity. What we also notice is that our contemporaneous effect size in innovation quality at a cross-country level is quite small relative to the findings by Ernst et al. (2013). This can be attributed to the fact we use both granted and non-granted patents for our micro-aggregate, whilst Ernst et al. (2013) only used granted-patents. We may ascribe the statistical insignificance in column 1 and the small contemporaneous impact size to the effect of corporate tax on innovation activity not occurring instantaneously.

To test this, we use our specification from equation (1) by replacing the contemporaneous EATR in columns 1 and 4 with a three year lag to take account of the inertia of innovation. In columns 2 and 5, we estimate a simple PPML regression of innovation quantity and innovation quality on a three year lag of EATR. In both regressions, the effects are statistically significant at the 5% level, implying the inverse effects of EATR on innovation are sluggish. The effect of a three year lag of EATR on innovation output is greater in absolute terms than the contemporaneous EATR, justifying that the response of innovation activity takes time to occur. But aggregate variables are omitted in columns 2 and 5, which creates a greater effect in absolute terms for the coefficient of EATR. Note we lost 78 observations because we are using a lag of three years and thus each country loses three observations.[9]

In the remaining columns, column 3 and column 6, we introduce cross-country controls to extend our simple PPML regressions. Contrary to Bosenberg and Egger (2017) who find a statistically negative effect of GDP Per Capita on innovation activity this year, we find that GDP Per Capita has a positive and significant effect on both innovation activity and quality of innovation. Again, we may attribute this to a larger cross-section: Bosenberg and Egger (2017) include 106 economies containing both developed and developing countries, whilst we only use 26 developed economies. One possible explanation is that the regressions in Bosenberg and Egger (2017) gives more weight to developing countries: agents living in developing countries may use income earned on basic needs such as food and housing rather than innovation and thus innovation falls.

R&D to GDP ratio is positive and statistically significant for innovation quantity and innovation quality. This explains that as governments increase their R&D expenditures this year, we can expect innovation output to increase in the same year. The coefficient for population growth is negative and is statistically significant for both dependent variables. This result is consistent with

[9]We have 26 countries: $26 \times 3 = 78$.

Coccia (2014) who finds the population growth has an inverse relationship with innovation in the contemporaneous year. Openness is negative and marginally significant for innovation activity but statistically insignificant for innovation quality. The statistical insignificance of openness and its negative coefficient can be ascribed to their effect occurring in the same period: the effects are sluggish. This may also be the case with the variable college, where it is statistically insignificant for both dependent variables.

The number of researchers working in R&D is negative and significant. Our results differ to Shakhmuradyan (2022) who find a positive relationship in the contemporaneous effect of the number of researchers and innovation. For our results, we consider the variation in the country and time specific effects to be correlated with the independent variables. As a result, we control for these factors by introducing controls for fixed effects in our model. Shakhmuradyan (2022) on the contrary, uses a OLS random effects model, which assumes the individual specific effects are uncorrelated with the independent variables. Zero correlation between individual specific effects and independent variables may not necessarily be true and thus yield results different to our PPML.

With respect to our variable of interest in columns 3 and 6, they are both statistically significant at the 5% level and have a negative effect on innovation. Relative to the simple regression, the inclusion of cross-country controls have slightly reduced the effect of EATR on innovation activity and the quality of innovation in absolute terms.

Similar to the literature, we find that the effects of EATR on innovation quality has a negative effect, both contemporaneously and three years later. However, we find that the effects of EATR on a country's innovation activity is statistically insignificant in the contemporaneous year, but has a statistically significant effect in the third year. We test the robustness of our baseline results in the following subsection.

## 4.2 Robustness Checks

### 4.2.1 Alternative Dynamics

We have discussed the importance of innovation containing inertial behaviours. This was shown empirically in column 1 of table 2 where the contemporaneous effect of EATR on innovation quantity was statistically insignificant. But when we replaced the contemporaneous EATR with a three year lag, we found the effect was statistically significant at the 5% level.

However, by following Akcigit et al. (2022), Mukherjee et al. (2017), and Atanassov and Liu (2020) with a three year lag, this lag length seems quite arbitrary: does the effect of EATR on innovation apply for alternative dynamics? Table 3 relaxes the assumption that innovation is effected three years later by using alternative lag lengths for EATR. The results from table 3 follows the approach from columns 3 and 6 of table 2 but with an alternative lag length for EATR. Table 3 has derived particularly interesting results: the statistically significant negative effect of EATR on innovation quantity begins three years later, whilst EATR has no statistically significant effect on the quality of innovation after the third year. Note also that as the yearly

Table 3: Alternative Dynamics

| | Quantity$_t$ | | | | | Quality$_t$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| EATR$_{t-1}$ | -0.032 | | | | | -0.035*** | | | | |
| | (0.066) | | | | | (0.006) | | | | |
| EATR$_{t-2}$ | | -0.058 | | | | | -0.053** | | | |
| | | (0.054) | | | | | (0.024) | | | |
| EATR$_{t-3}$ | | | -0.083** | | | | | -0.064** | | |
| | | | (0.041) | | | | | (0.031) | | |
| EATR$_{t-4}$ | | | | -0.097*** | | | | | -0.042 | |
| | | | | (0.034) | | | | | (0.047) | |
| EATR$_{t-5}$ | | | | | -0.074** | | | | | 0.061 |
| | | | | | (0.031) | | | | | (0.043) |
| Observations | 442 | 416 | 390 | 364 | 338 | 442 | 416 | 390 | 364 | 338 |

Robust standard errors in parentheses. All specifications include country and year fixed effects. Contemporaneous cross-country variables are included in all specifications. Robust standard errors clustered at country level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

EATR lag increases, we lose more observations. This is due to the loss of each time series $t$ from lagging the EATR.

Although this thesis does not focus on multinationals and the shifting of intangible assets, it was discussed from the literature review that multinationals have a larger contribution toward a country's innovation output than domestic firms. Therefore, our results are consistent with Griffith et al. (2013) and Karkinsky and Riedel (2012) in that innovation activity for the host country falls as EATR rises due to multinationals shifting their inventions to alternative destination countries for tax benefit purposes. According to the management literature, multinationals tend to be larger in size, which implies higher levels of rigidity (see Griener (1998) amongst others). As a result, the sluggishness in the unobservable firm characteristics such as the negotiation and planning of patent relocation may be greater for multinationals. Thus, from a cross-country setting, we only observe the statistically significant fall in the number of patents filed three years later on average.

Multinationals also represent the majority of the quality of innovation for a given country (Almodovar & Nguyen, 2022). Therefore, one possible explanation for the fall in innovation quality is due to the patents of multinationals leaving the host country from an increase in the EATR. As the main concern for multinationals now is the relocation of their patents, effort towards creating high quality inventions have fallen and therefore the quality of innovation output has also declined. This effect is propagated up to the third year. Afterwards, we may now have only domestic firms who, as discussed, have marginal contribution toward innovation for a country. This is because domestic firms tend to be more financially constrained and would therefore be inelastic to changes in corporate tax rates (Cheung et al., 2021). Although we do not have data on whether firms are domestic or multinationals, we can theorize that as domestic firms have invariably low quality of innovation, EATR does not covary with domestic firms' innovation quality. Thus, the effect of an increase in EATR this period has no statistically significant effect on innovation quality when only domestic firms are present, or four years later.

### 4.2.2 Alternative Aggregation: Star Inventors

From the literature review, innovation activity and the quality of innovation for countries will fall given an increase in corporate tax rates as firms shift their profits to another region or relocate their intangible assets. These findings however, lacked the heterogeneity across the behaviour of the star inventors and non-star inventors. Star inventors are economic agents of paramount significance: they tend to be highly educated and highly productive for the economy. As a consequence, the presence of these top inventors attract more multinational firms.

With data on 260,000 star US inventors from 1977 to 2010, Moretti and Wilson (2017) found that the long-run elasticity of firm mobility relative to the effective average corporate tax rate is 1.9

Similarily, Akcigit, Baslandze and Stantcheva (2016) examined how personal income tax impacts inventor's choice of location. They collected patent data on the US Patent and Trademark Office from 1977 to 2000, focusing on 8 OECD countries that represent the bulk of countries the inventors are from. They found that star inventors are more elastic to changes in personal income tax rates than non-star inventors. The elasticities are even greater if star inventor works for a multinational company.

We attempt to adopt this framework by Akcigit et al. (2016) and Moretti and Wilson (2017) to a cross-country setting. To do this, we dissect our sample into two subsamples. In one subsample, we aggregate only star inventors. In the other subsample is a micro-aggregate of non-star inventors. Table 4 builds on columns 3 and 6 from table 2 by presenting the difference in innovation relative to EATR for micro-aggregates with solely star inventors and micro-aggregates of solely non-star inventors.

Table 4: Alternative Aggregation: Star and Non-Star Inventors

| | Top 5% | | Bottom 95% | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Quantity$_t$ | Quality$_t$ | Quantity$_t$ | Quality$_t$ |
| EATR$_{t-3}$ | -0.154*** | -0.113** | -0.064 | -0.075* |
| | (0.044) | (0.057) | (0.045) | (0.043) |
| GDPPerCap$_t$ | 1.368*** | 1.564*** | 0.681** | 1.292*** |
| | (0.357) | (0.432) | (0.306) | (0.292) |
| R&D-GDP$_t$ | 0.789*** | 0.885*** | 0.605*** | 0.765*** |
| | (0.283) | (0.259) | (0.227) | (0.247) |
| PopulationGrowth$_t$ | 0.026 | 0.022 | -0.174*** | -0.192*** |
| | (0.017) | (0.017) | (0.055) | (0.051) |
| Openness$_t$ | -0.785* | -0.545 | -0.377 | -0.747*** |
| | (0.401) | (0.407) | (0.233) | (0.252) |
| Researchers$_t$ | -0.542*** | -0.471** | -0.311** | -0.514*** |
| | (0.201) | (0.232) | (0.125) | (0.137) |
| College$_t$ | 0.435** | 0.217 | 0.128 | 0.276* |
| | (0.190) | (0.238) | (0.135) | (0.143) |
| Observations | 335 | 335 | 390 | 390 |

Robust standard errors in parentheses.
All specifications include country and year fixed effects.
Robust standard errors clustered at country level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Divided into two blocks, the left block contains columns 1 and 2 which uses only the aggregation of star inventors; whilst the right block is the subsample which aggregates non-star inventors. The number of observations differ across the two subsamples. This is because some countries at a given year may not have star inventors.

Beginning with innovation quantity as our dependent variable, we see that the elasticity is greater in absolute value when aggregates only consists of star inventors rather than non-star inventors. Moreover, the effect of a 1% increase in EATR three years ago has a statistically significant effect for countries with solely star inventors, whilst its effect is not statistically significant for countries without star inventors.

For the quality of innovation, again, countries with star inventors are more elastic to an increase in EATR relative to countries who do not have star inventors. The coefficient for EATR in column 2 is statistically significant at the 5% level, whilst it is marginally significant for column 4.

Some speculations are applicable here. Akcigit et al. (2016) and Moretti and Wilson (2017) uses star inventor location as their dependent variable and found that the propensity to migrate increases as corporate and personal tax increases. With a fall in the number of star inventors, they suspect innovation levels to also fall. Our regressions are consistent with the results by Akcigit et al. (2016) and Moretti and Wilson (2017): as the number of star innovators fall, innovation output will also fall. For countries with only star inventors, firms will want to shift their patents to an alternative region for tax benefits. As a result, innovation activity for a country given an increase in the EATR will fall. Furthermore, multinationals dominate the quality of innovation and thus the fall in innovation quality for a country, supporting the view that star inventors migrate to another location. This can be justified by the statistical significance at the 10% level in column 4, where the bottom 95% of citations could be mainly domestic firms, whose contribution toward the quality of innovation for a given country is marginal.

### 4.2.3 Alternative Aggregation: Technology Classes

Patent intensity can differ across technology classes. Subsequently the innovation quality and its rate of obsolescence across technology classes also varies. In this subsection, we measure the effects of EATR on innovation by taking account of the heterogeneity across technology classes. This was achieved by micro-aggregating patents conditional on the first digit of a patent's CPC code. This alternative micro-aggregation creates countries that only patents one specific technology class. As a result, this exposes us to a country's technology class specialization. We could, of course, aggregate patents based on their four digit CPC to analyse the heterogeneity across technology classes even further, but that would be beyond the scope of this thesis. Therefore, each column represents countries that contain patents based on the first digit of their CPC. Table 5 and 6 extends the PPML regression from column 3 and 6 of table 2 to a cross-country-CPC setting, respectively.

First note that for both tables 5 and 6, the number of observations differ across each technology class regression. This is because not all countries have a patent in that technology class at a given year.

For the majority of technology classes, our results in table 5 are consistent with our baseline results: a 1% increase in EATR three years ago reduces innovation activity this year. All technology classes are statistically significant at the 5% level, with the exception of the Textile, Physics and Electricity sectors. Though the statistical insignificance of the Textile technology class can be justified to its lower number of observations and patent intensity relative to other technology classes, the insignificance of the Physics and Electricity technology classes is peculiar given they are patent-heavy technology classes. Despite the purpose of this subsection is to examine the heterogeneity across the first digit of the CPC code, if we account for the first four digits of the CPC code, we may derive statistically significant results for the Physics and Electricity technology classes.

We now move onto table 6, where innovation quality is our dependent variable. With the exception of the Human Necessities, Textile, Fixed Construction and Electricity technology classes, our variable of interest is statistically significant at the 5% level. However, the Physics technology class is positive and significant: if EATR were to increase by 1% three years ago, the quality of

Table 5: Cross-CPC Regressions: Innovation Quantity

| | (1) A Human Necessities | (2) B Transport | (3) C Chemistry | (4) D Textile | (5) E Fixed Construction | (6) F Mechanical Engineering | (7) G Physics | (8) H Electricity | (9) Y General |
|---|---|---|---|---|---|---|---|---|---|
| $\text{EATR}_{t-3}$ | -0.194*** | -0.143*** | -0.086*** | 0.051 | -0.206*** | -0.256*** | 0.079 | 0.072 | -0.161** |
| | (0.020) | (0.045) | (0.024) | (0.081) | (0.061) | (0.097) | (0.069) | (0.107) | (0.075) |
| $\text{GDPPerCap}_t$ | 0.533*** | 0.791** | 0.825*** | -1.361** | 0.586* | 0.586 | 1.291*** | 1.219** | 0.825* |
| | (0.091) | (0.327) | (0.115) | (0.563) | (0.341) | (0.398) | (0.104) | (0.512) | (0.464) |
| $\text{R\&D-GDP}_t$ | 0.434*** | 0.353 | 0.426*** | 0.496 | 0.673*** | 0.702*** | 1.012*** | 0.443 | 0.442 |
| | (0.067) | (0.299) | (0.085) | (0.320) | (0.245) | (0.257) | (0.076) | (0.339) | (0.389) |
| $\text{PopulationGrowth}_t$ | -0.120*** | -0.138** | -0.119*** | 0.147 | 0.085 | -0.174** | -0.283*** | -0.224* | -0.234*** |
| | (0.018) | (0.056) | (0.023) | (0.097) | (0.078) | (0.074) | (0.021) | (0.116) | (0.082) |
| $\text{Openness}_t$ | 0.033 | -0.325 | -0.531*** | -1.445*** | -0.649* | -1.264*** | -0.535*** | -0.525 | -1.331*** |
| | (0.096) | (0.278) | (0.113) | (0.274) | (0.358) | (0.397) | (0.101) | (0.555) | (0.346) |
| $\text{Researchers}_t$ | 0.030 | -0.361*** | -0.288*** | 0.069 | -0.194 | -0.183 | -0.803*** | -0.242 | -0.252 |
| | (0.055) | (0.124) | (0.072) | (0.276) | (0.221) | (0.152) | (0.060) | (0.322) | (0.256) |
| $\text{College}_t$ | 0.282*** | 0.199 | 0.296*** | 0.747*** | 0.246** | 0.436** | 0.232*** | -0.122 | 0.144 |
| | (0.052) | (0.131) | (0.064) | (0.218) | (0.101) | (0.185) | (0.056) | (0.318) | (0.269) |
| $N$ | 355 | 354 | 362 | 240 | 307 | 334 | 351 | 348 | 345 |

Robust standard errors in parentheses. All specifications include country and year fixed effects.

Robust standard errors clustered at country level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Cross-CPC Regressions: Innovation Quality

| | (1)<br>A<br>Human Necessities | (2)<br>B<br>Transport | (3)<br>C<br>Chemistry | (4)<br>D<br>Textile | (5)<br>E<br>Fixed Construction | (6)<br>F<br>Mechanical Engineering | (7)<br>G<br>Physics | (8)<br>H<br>Electricity | (9)<br>Y<br>General |
|---|---|---|---|---|---|---|---|---|---|
| $\text{EATR}_{t-3}$ | -0.038 | -0.164** | -0.073*** | 0.308 | -0.065 | -0.205** | 0.110*** | -0.061 | -0.186** |
| | (0.030) | (0.078) | (0.023) | (0.210) | (0.138) | (0.103) | (0.023) | (0.136) | (0.090) |
| $\text{GDPPerCap}_t$ | 1.308*** | 1.562*** | 2.348*** | -2.069** | 1.369** | 1.379*** | 2.349*** | 1.748* | 1.089 |
| | (0.104) | (0.458) | (0.140) | (0.804) | (0.621) | (0.455) | (0.116) | (0.908) | (0.749) |
| $\text{R\&D-GDP}_t$ | 1.091*** | 0.338 | 2.040*** | 0.678 | 0.650 | 1.271*** | 0.703*** | -0.068 | 1.328*** |
| | (0.082) | (0.481) | (0.097) | (1.056) | (0.579) | (0.471) | (0.089) | (0.601) | (0.462) |
| $\text{PopulationGrowth}_t$ | 0.009 | -0.059 | -0.225*** | 0.179 | 0.083 | -0.041 | -0.364*** | -0.391*** | -0.092 |
| | (0.020) | (0.065) | (0.025) | (0.157) | (0.112) | (0.109) | (0.023) | (0.105) | (0.147) |
| $\text{Openness}_t$ | 0.712*** | -0.292 | -0.823*** | -1.125 | -0.289 | -1.463*** | -0.710*** | 0.109 | -0.961** |
| | (0.093) | (0.517) | (0.114) | (0.899) | (0.611) | (0.524) | (0.103) | (0.683) | (0.448) |
| $\text{Researchers}_t$ | -0.514*** | -0.401* | -0.620*** | -0.692 | -0.147 | -0.366 | -0.972*** | -0.513 | -0.020 |
| | (0.066) | (0.233) | (0.076) | (0.668) | (0.480) | (0.309) | (0.069) | (0.378) | (0.418) |
| $\text{College}_t$ | 0.030 | 0.132 | -0.722*** | 0.727 | 0.660** | 0.654* | 0.119* | 0.593** | 0.232 |
| | (0.067) | (0.306) | (0.072) | (0.575) | (0.304) | (0.353) | (0.068) | (0.249) | (0.320) |
| $N$ | 355 | 354 | 362 | 240 | 307 | 334 | 351 | 348 | 345 |

Robust standard errors in parentheses. All specifications include country and year fixed effects.

Robust standard errors clustered at country level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

innovation for countries only consisting of patents from the Physics technology class would rise by 0.11% this period. Although this is a peculiar result, we may speculate that as tax revenues increase from corporate tax raises, governments apply greater efforts on innovation in the Physics technology class through extra finances. This is, however, a conjecture.

Along with this, Human Necessities is not statistically significant. Amongst the subsets of Human Necessities are the Pharmaceutical and Baking technology subclasses, which are patent-heavy and less-patent-heavy technology classes, respectively. As a result, similar to the Electronic technology class, this statistical insignificance can be attributed to taking account of solely the first digit of the four digit CPC code.

All in all, this section has explored and taken account of the heterogeneity across technology classes on a 1% increase in EATR three years ago and its effect on innovation output this year. For the most part, our results are consistent with our baseline findings and the literature. However, we came across a few peculiar results, which can be examined in another research paper.

#### 4.2.4 Truncation Bias

The final robustness check we will implement is based on our sample's truncation bias problem. To remind ourselves, truncation bias is when patents filed toward the end of the sample have less time to accumulate cites and may appear of lower quality. This implies that the high quality patents tend to be bunched toward the beginning of the sample. Our control variables will therefore be biased toward zero due to a lack of variation in innovation quality toward the end of the sample.

Truncation bias in our dataset was illustrated in Figure 1 in the Patent Data section of this thesis. Moreover, Figure 1 showed that the micro-aggregate of non-star inventors had more cites than the star inventors around 2013, though this is merely attributed to the aggregation of higher quantities of patents with lower cites.

From this, despite controlling for year dummies, the EATR may be effected from the end of sample truncation bias. To test the robustness, we truncate the ending sample at 2013 as this was approximately when the data illustrated quantity of patents had more value than the forward cites it has. Although choosing the year 2013 for truncation is somewhat arbitrary, we include truncating our sample at 2014 to provide a more robust conclusion. Table 7 is an extension of columns 3 and 6 from table 2 by accounting for truncation bias. Table 7 presents our results on ending our sample at earlier years of 2013 and 2014.

Table 7: Early Ending Samples

| | Sample Ending 2014 | | Sample Ending 2013 | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Quantity$_t$ | Quality$_t$ | Quantity$_t$ | Quality$_t$ |
| EATR$_{t-3}$ | -0.026** | -0.065** | -0.014** | -0.060** |
| | (0.013) | (0.028) | (0.006) | (0.025) |
| GDPPerCap$_t$ | 0.761*** | 1.608*** | 0.662** | 1.595*** |
| | (0.280) | (0.322) | (0.266) | (0.330) |
| R&D-GDP$_t$ | 0.508** | 0.797*** | 0.656*** | 0.880*** |
| | (0.246) | (0.261) | (0.213) | (0.259) |
| PopulationGrowth$_t$ | -0.144*** | -0.127** | -0.120*** | -0.113* |
| | (0.048) | (0.057) | (0.046) | (0.060) |
| Openness$_t$ | -0.492*** | -0.363 | -0.504*** | -0.343 |
| | (0.189) | (0.310) | (0.166) | (0.307) |
| Researcher$_t$ | -0.354*** | -0.490** | -0.443*** | -0.510** |
| | (0.117) | (0.194) | (0.167) | (0.221) |
| College$_t$ | 0.268** | 0.219 | 0.214* | 0.212 |
| | (0.133) | (0.204) | (0.128) | (0.214) |
| Observations | 312 | 312 | 286 | 286 |

Robust standard errors in parentheses.
All specifications include country and year fixed effects.
Robust standard errors clustered at country level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

From table 7, despite tackling truncation bias, our results corroborate our baseline result of a negative and statistically significant effect of a 1% increase in EATR three years ago on innovation output this year.

# 5 Conclusion

The purpose of this thesis was to examine the relationship between corporate tax and innovation at a cross-country setting whilst accounting for innovation's inertial behaviours. We factor in the sluggishness of innovation by simply replacing the literature's commonly interpreted contemporaneous corporate tax variable with a three year lag, as implemented by Akcigit et al. (2022), Mukherjee et al. (2017), and Atanassov and Liu (2020).

We first test whether the inverse simultaneous relationship at a cross-country setting is in line with the findings from the literature. With patent quality as our measurement for innovation quality, we found that a 1% increase in the EATR has a statistically significant negative effect on a country's innovation quality in the same year, which aligns to the findings by Ersnt et al.

(2013). However, unlike Bosenberg and Egger (2017) who find a significant inverse relationship between EATR and innovation quantity in the same year, we found that an increase in the EATR had a zero effect on innovation activity for any given country. We speculate that this statistical insignificance is attributed to the contemporaneous effect: firms may not have time to respond. As a consequence, when we replaced the contemporaneous variable with a three year lag, we found that the effects of EATR on both innovation activity and the quality of innovation to be statistically significant at the 5% level.

The three year lag may be somewhat arbitrary. To tackle this, we introduce alternative lag lengths: yearly lags of one to five. We found that a 1% increase in EATR has no statistically significant effect on a country's innovation activity until three years later. Griffith et al. (2013) amongst others argue a fall for a country's innovation activity given an increase in corporate tax can be attributed to multinationals shifting their profits. We therefore speculate that the delayed significant effect is attributed to unobservable characteristics of multinationals, such as the negotiation and planning of the shifting of intangible assets.

With respect to the quality of innovation, our findings show that a 1% increase in EATR three years ago has a statistically significant effect for a country's innovation quality for up to three years. The effects four and five years later are statistically insignificant. We construct a theory that follows the findings of Almodovar and Nguyen (2022): foreign subsidiaries has a strong influence on a country's innovation levels whilst domestic firms' contribution are marginal. From this, we therefore argue that as multinationals leave in the fourth year, only domestic firms are present for any given country. As domestic firms have invariably little contribution toward a country's innovation quality, its covariance with EATR is not statistically significant.

We then examined the elasticity of countries with solely star inventors and solely non-star inventors. Consistent with Akcigit et al. (2016) and Moretti and Wilson (2017), we found that the innovation output of micro-aggregates with only star inventors tend to have a higher elasticity in absolute terms relative to micro-aggregates with solely non-star inventors. Along with this, we found that the effects of EATR on innovation for countries with solely star inventors are statistically significant, whilst its effect for countries with only non-star inventors is not statistically significant at the 5% level. Again, this may be due to non-star inventors working for domestic firms, whilst star inventors work for multinationals.

Apart from the heterogeneity across the quality of inventors, we also dissected our micro-aggregate across technology classes based on the first digit of the four digit CPC code. We found across most technology classes that a 1% increase in EATR three years ago reduces innovation levels this years. However, there were some technology classes which were statistically insignificant. There are two possible reasons for this.

The first reason is that we only used the first digit of the CPC code to take account of heterogeneity across technology classes: the first digit contains a vast amount of subclasses. As a result, were one interested in digging deeper into the heterogeneity across technology classes, they would use the four digit CPC code which may derive statistically significant results. Secondly, the patent intensity and thus patent quality differs across technology classes. Consequently, the

EATR may not necessarily have a statistically significant effect on innovation for these technology classes.

Lastly, we took account of a common problem when measuring innovation quality: truncation bias. We mitigate truncation bias by cutting our time series by three to four years from 2017. Our findings here nevertheless corroborate our baseline results of a negative relationship between a three year lag of EATR and innovation output this year.

Though this thesis helped contribute in understanding the lag effect of EATR on innovation output at a cross-country setting, this thesis contain certain limits that would be of paramount importance in the academic literature. First, we theorized from the academic literature that star inventors and innovation output are mainly attributed to multinational firms. Unfortunately, we do not have data to test this. Therefore, it would be worthwhile to collect firm-level data on star inventors and test the difference in innovation output with respect to whether they belong to domestic firms or multinationals.

Secondly, although literature investigating the Human Necessities technology class from the CPC is abundant, there remains scant evidence on the relationship between corporate tax and Human Necessities and other specific technology classes. This will thus elucidate the literature on the elasticity of innovation output from changes in corporate taxes across idiosyncratic technology classes.

Lastly, it is worth investigating knowledge spillover of multinationals who shift their patents to an alternative country along with the knowledge retained within the host country. This will allow us to further understand the ability of domestic firms in diffusing technical foreign knowledge.

Thus, the limitations from this thesis can greatly elucidate the field of innovation and economic growth.

# 6 References


1. Acemoglu, D. 2002. "Directed Technical Change." *Review of Economic Studies.* 69(4). Pp 781 - 809.

2. Acemoglu, D. & Linn, J. 2004. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry." *Quarterly Journal of Economics.* 119(3). Pp 1049-1090.

3. Aghion, P., Dechezleprêtre, A., Hémous, D., Martin, R., & Van Reenen, J. 2016. "Carbon Taxes, Path Dependency, and Directed Technical Change: Evidence from the Auto Industry ." *Journal of Political Economy.* 124(1). Pp 1 - 51.

4. Aghion, P. & Howitt, P. 1992. "A Model of Growth Through Creative Destruction." *Econometrica.* 60(2). Pp 323 - 351.

5. Akcigit, U., Baslandze, S. & Stancheva, S. 2016. "Taxation and the International Mobility of Inventors". *American Economic Review.* 106(10). Pp 2930 - 2981.

6. Akcigit, U., Grigsby, J., Nicholas, T. & Stancheva, S. 2022. "Taxation and Innovation in the Twentieth Century". *Quarterly Journal of Economics.* 137(1). Pp 329 - 385.

7. Almodovar, P. and Nguyen, Q. 2022. "Product innovation of domestic firms versus foreign MNE subsidiaries: The role of external knowledge sources." *Technological Forecasting and Social Change.* Vol 184.

8. Alstadsæter, A., Barrios, S., Nicodeme, G., Skonieczna, A.M. & Vezzani, A. 2018. "Patent boxes design, patents location, and local R&D". *Economic Policy.* 33(93). Pp 131 - 177.

9. Atanassov, J. & Liu, X. 2020. "Can Corporate Income Tax Cuts Stimulate Innovation?" *Journal of Quantitative and Financial Analysis.* 55(5). Pp. 1415–1465.

10. Bloom, N., Griffith, R. & Van Reenan, J. 2003. "Do R&D tax credits work? Evidence from a panel of countries 1979–1997." *Journal of Public Economics.* Vol 85. Pp 1 - 31.

11. Bosenberg, S. & Egger, P. 2017. "R&D tax incentives and the emergence and trade of ideas". *Economic Policy.* 32(89). Pp 39 - 80.

12. Cai, J., Chen, Y. & Wang, X. 2018. "The Impact of Corporate Taxes on Firm Innovation: Evidence from the Corporate Tax Collection Reform in China". *National Bureau of Economic Research Working Paper.* Working Paper 25146.

13. Chen, W. & Miller, K. 2007. "Situational and institutional determinants of firms' R&D search intensity". *Strategic Management Journal.* 28(4). Pp 369-381.

14. Chen, Z., Liu, Z., Serrato, J.C.S., & Xu, D.Y. 2021. "Notching R&D Investment with Corporate Income Tax Cuts in China". *American Economic Review.* 111(7). Pp 2065 - 2100.

15. Cheung, A., Guo, P., Weng, C. & Wu, Q. 2021. "Innovation and Corporate Tax Planning: The Distinct Effects of Patents and R&D". *Contemporary Accounting Research.* 18(1). Pp 621 - 653.

16. Coccia, M. 2014. "Driving forces of technological change: The relation between population growth and technological innovation: Analysis of the optimal interaction across countries." *Technological Forecasting and Social Change.* Vol 82. Pp 52 - 65.

17. Costinot, A., Donaldson, D., Kyle, M., & Williams,H. 2019. "The More We Die, The More We Sell? A Simple Test of the Home-Market Effect." *Quarterly Journal of Economics.* 134(2). Pp 843–894.

18. Davies, R.B., Kogler, D.F. & Hynes, R. 2020. "Patent Boxes and the Success Rate of Applications." *CESifo Working Paper No. 8375.*

19. Djankov, S., Ganser, T., McLiesh, C., Ramalho, R. & Shleifer, A. 2011. "The Effect of Corporate Taxes on Investment and Entrepreneurship". *American Economic Journal: Macroeconomics* 2(3). Pp 31–64.

20. Ernst, C., Richeter, K. & Riedel, N. 2013. "Corporate Taxation and the Quality of Research and Development." *ZEW Discussion Paper No. 13-010.*

21. European Innovation Scoreboard 2023 *European Innovation Scoreboard.* Accessed on 25 July 2023. Available from:
https://research-and-innovation.ec.europa.eu/statistics/performance-indicators/european-innovation-scoreboard_en

22. Gentry, W. & Hubbard, G. 2005. "'Success Taxes,' Entrepreneurial Entry, and Innovation." *Innovation Policy and the Economy.* Vol 5. Pp 87 - 108.

23. Goolsbee, A. 1998. "Does Government R&D Policy Mainly Benefit Scientists and Engineers?" *American Economic Review.* 88(2). Pp 298 - 302.

24. Griener, L. 1998. "Evolution and Revolution as Organizations Grow." *Harvard Business Review.* Accessed on 12 August 2023. Sourced from:
https://hbr.org/1998/05/evolution-and-revolution-as-organizations-grow

25. Griffith, R., Miller, H. & O'Connell, M. 2014. "Ownership of intellectual property and corporate taxation". *Journal of Public Economics.* Vol 114. Pp 12 - 23.

26. Grossman, G. & Helpmann, E. 1990. "Trade, Innovation, and Growth." *American Economic Review.* 80(2). Pp 86 - 91.

27. Grubert, H. & Slemrod, J. 1998. "The Effect Of Taxes On Investment And Income Shifting To Puerto Rico." *The Review of Economics and Statistics.* 80(3). Pp 365 - 373.

28. Hall, B., Jaffe, A. & Trajtenberg, M. 2001. "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools." *National Bureau of Economic Research Working Paper.* Working Paper 8498.

29. Hall, B. & Van Reenen, J. 2009. "How Effective are Fiscal Incentives for R&D? A review of the Evidence," *Research Policy.* 29(4-5). Pp 449-469.

30. Hanlon, W. 2015. "Necessity Is the Mother of Invention: Input Supplies and Directed Technical Change." *Econometrica.* 83(1). Pp 67-100.

31. He, L., Jiang, X. & Fang. L. 2023. "Tax policy reform and corporate innovation in China". *Finance Research Letters.* Vol 55.

32. Howell, A. 2016. "Firm R&D, innovation and easing financial constraints in China: Does corporate tax reform matter?" *Research Policy.* 45(10). Pp 1996 - 2007.

33. Hsu, P., Tian, X. & Xu, Y. 2014. "Financial Development and Innovation: Cross-country Evidence." *Journal of Financial Economics.* 112(1). Pp 116-135.

34. Karkinsky, T. & Riedel, N. 2012. "Corporate Taxation and the choice of patent location within Multinational Firms." *Journal of International Economics.* 88(1). Pp 176 - 185.

35. Krieger, J. 2021. "Trials and Terminations: Learning from Competitors' R&D Failures." *Management Science.* Vol 67. No 9.

36. Lerner, J. & Seru, A. 2022. "The Use and Misuse of Patent Data: Issues for Finance and Beyond". *Review of Financial Studies.* 35(6). Pp 2667 - 2704.

37. Linder, S.B. 1961. *An essay on trade and transformation.* Pp. 82-109. Stockholm: Almqvist & Wiksell.

38. Manso, G. 2011. "Motivating innovation," *Journal of Finance.* 66(5). Pp 182-1860.

39. Moretti, E. & Wilson, D. 2017. "The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists". *American Economic Review.* 107(7). Pp 1858–1903.

40. Moscona , J. & Sastry, K. 2023. "Does Directed Innovation Mitigate Climate Damage? Evidence from U.S. Agriculture." *Quarterly Journal of Economics.* 138(2). Pp 637 - 701.

41. Mukherjee, A., Singh, M. & Žaldokas, A. 2017. "Do corporate taxes hinder innovation?" *Journal of Financial Economics.* 24(1). Pp 195 - 221.

42. Packalen, M. & Bhattacharya, J. (2019) "Age and the Trying Out of New Ideas". *Journal of Human Capital.* 13(2). Pp 341-373.

43. Romer, C. & Romer, D. 2010. "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks". *American Economic Review.* 100(3). Pp 763-801.

44. Romer, P. 1990. "Endogenous Technological Change". *Journal of Political Economy.* 98(5). Pp S71-S102.

45. Shakhmuradyan, G. 2022. "Does Fiscal Policy matter for business R&D investment? Panel Data evidence from Central and Eastern Europe." *Central European Business Review.* 11(3). Pp 79 - 96.

46. Tian, X & Wang T.Y. 2011. "Tolerance for Failure and Corporate Innovation," *Review of Financial Studies.* 27(1). Pp 211-255.

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

48. Wang, E. 2010. "Determinants of R&D Investment: The Extreme-Bounds-Analysis Approach Applied to 26 OECD Countries". *Research Policy.* 39(1). Pp 103–116.

# 7 Appendix

## 7.1 Appendix A: Cooperative Patent Classification Explanation

In this section of the appendix, we will briefly explain EPO's CPC system. The purpose of the CPC is to assign each invention to a technology class, or technology classes in some cases. We can almost view technology classes as synonymous to sectors or industries but within the realm of patents. Table A1 presents the first digit CPC codes.

Table A1: First Digit CPC Code

| First CPC Digit | CPC section |
|---|---|
| A | Human Necessities |
| B | Performing Operations and Transport |
| C | Chemistry |
| D | Textile and Paper |
| E | Fixed Construction |
| F | Mechanical Engineering |
| G | Physics |
| H | Electricity |
| Y | General |

As we can see, the first digit of the CPC code contains quite diverse technology classes. Although rare in its appearance, the technology class General is applied for patents whose invention does not fall within A through H. Therefore, this patent can be considered an emerging technology class, or general.

The first digit CPC is quite broad. Therefore, EPO's CPC system tend to have subclasses, which is a four digit code in our dataset. The first digit is a letter from Table A1; the second and third digit are numbers from 0 to 9; and the fourth digit is again a letter from all 26 letters of the alphabet.[10] For example, Human Necessities contains subclasses that can range from pharmaceutical inventions to home baking inventions. A patent may also contain more than one technology class. This represents the brevity of the invention: the invention can be used in other technology classes. The table below shows one patent from our initial dataset with multiple subclasses.

Table A2: Example of CPC subclasses from one patent

| CPC | CPC section |
|---|---|
| B01J | Chemical or Physical Processes |
| C40B | Combinatorial Chemistry |
| G01N | Determining Physical or Chemical Properties |

[10]This implies there are $9 \times 10 \times 10 \times 26 = 234,000$ possible patent subclasses!

This patent has multiple CPCs because its innovation can contribute to many strands of inventions. Thus, using a four digit CPC code allow researchers to identify and understand certain inventions and technology classes more easily.

## 7.2 Appendix B: Country-CPC Weight Example

We provide an explicit example on the methodology for a cross-country-CPC weight for one invention in table A3.

| Country | Inventor ID | CPC | Country Weight | CPC Weight | Country-CPC Weight |
|---|---|---|---|---|---|
| Ireland | 1 | A | 0.67 | 0.5 | 0.335 |
| Ireland | 1 | B | 0.67 | 0.5 | 0.335 |
| Ireland | 2 | A | 0.67 | 0.5 | 0.335 |
| Ireland | 2 | B | 0.67 | 0.5 | 0.335 |
| Germany | 3 | A | 0.33 | 0.5 | 0.165 |
| Germany | 3 | B | 0.33 | 0.5 | 0.165 |

| Country | Inventor ID | CPC | Country Weight | CPC Weight | Country-CPC Weight |
|---|---|---|---|---|---|
| Ireland | 1 | A | 0.67 | 0.5 | 0.335 |
| Ireland | 1 | B | 0.67 | 0.5 | 0.335 |
| Germany | 3 | A | 0.33 | 0.5 | 0.165 |
| Germany | 3 | B | 0.33 | 0.5 | 0.165 |

Table A3: Example of Country-CPC weight methodology for one patent

In the top half of table A3, we have 3 inventors for one patent: two Irish and one German. This one patent has two CPCs: A and B.[11] We can view the distinct integers in variable Inventor ID as individual inventors. The first step is to fractionally weigh the patent by country. In this example, each Irish and German inventor would receive fractional weights of 0.67 and 0.33, respectively. If we were interested in solely weighing innovation levels with respect to country, we would keep one of each inventor from each country and the country weight ends here. If we did not drop duplicates with respect to country, we would be double counting this one patent.

For instance, each invention has a patent quantity of one. Dropping duplicates with respect to countries implies Ireland receives 0.67 from this one patent whilst Germany receives 0.33. The addition of 0.67 and 0.33 leads to a patent quantity of 1 for the one invention. If we did not drop the country duplicates, we would have patent quantity of $4 \times 0.67 \approx 2.68$ for Ireland and $2 \times 0.33 \approx 0.67$ for Germany, where the summation is well excess the patent quantity for one invention.

We also have two CPCs, A and B. As each inventor has contributed to this patent, every inventor contains these CPCs. Both A and B occurs once each, and therefore a fractional weight

[11]Note in our dataset we used all four CPC digits: we only present the first digit here for simplicity purposes.

of 0.5 is applied. So far, we have two individual fractional weights: a country weight and a CPC weight. We want to turn it into a Country-CPC weight, and so we multiply them together, as shown in column 6 in the top half of table A3.

Again, the aggregation across inventors for a given country for our Country-CPC weight in column 6 of the top half of table A3 is greater than one because of duplicates in the Irish inventor as they contain the same country and CPC characteristics. Therefore, we drop duplicates here with respect to country and CPC. This is shown in the bottom half of table 1: we have dropped one Irish inventor to derive our final Country-CPC weight, which sums to 1.[12] As these are weights, we can multiply this to each patent's quantity variable (which is one) and forward citations variable. Lastly, we aggregate the Country-CPC quantity and Country-CPC forward citations.

[12]Again, if we did not drop this one Irish inventor, our patent count for this one invention will be well excess of 1.