The causal effect of EU membership on innovation

A difference-in-difference approach

Markus Sageng Gyene

Thesis submitted for the degree of Master of Philosophy in Economics

Department of Economics

UNIVERSITY OF OSLO

May 2018
The causal effect of EU membership on innovation

A difference-in-difference approach

Markus Sageng Gyene
The causal effect of EU membership on innovation

http://www.duo.uio.no/

Printed: Reprosentralen, University of Oslo
Abstract

This thesis investigates the relationship between innovation and European Union membership using panel data on the firm level. I use a difference-in-difference estimator, considering access to the inner EU market as the treatment. This enables me to test the hypothesis that firms in countries that achieve membership status, change their innovation efforts by more or less compared to the other firms. By using fixed effects models I find that there is a significant percentage point decline in innovation efforts by firms in the new membership countries relative to the change for the control group firms. The effect of accession is present in the year of entry, but disappears in the the post-treatment years. Moreover, I find that the results are robust after controlling for average firm size.
Preface

Writing this thesis has been both inspiring and challenging. I am truly grateful to my supervisor, Professor Andreas Moxnes, for valuable guidance throughout the process. While I was writing this thesis I also worked as a research assistant at the Department of Economics. During that time they provided me with a work space, which made my days of writing easier. Also, I would like to thank Eurostat, who accepted my request for the data I have used.

Since this marks the end of a five years study, I would like to thank my friends. We have shared many good memories. Finally, big thanks to my parents that have always supported me throughout my time at the University.

I am solely responsible for any remaining errors in the thesis.

Oslo, May 2018

Markus Sageng Gyene
Contents

1 Introduction 1
2 Background 3
3 Economic mechanisms 7
  3.1 Theoretical framework 7
  3.2 Technology diffusion and knowledge spillovers 11
4 Data 14
  4.1 Survey data on innovation 14
  4.2 Descriptive statistics 15
5 Empirical analysis 20
  5.1 Methodology 20
  5.2 Results 21
  5.3 Validity of the study 28
  5.4 Discussion of the findings 29
6 Conclusion 32
7 Appendix 35
1 Introduction

The objective of this thesis is to investigate if increased openness of the economies through membership in the European Union motivate firms to increase their innovation. In particular, I want to test the hypothesis that firms change their innovation efforts as a result of being included in the EU inner market, compared to the change in innovation for firms that are not directly affected by the policy changes that applies to the new member countries.

Economic theory, such as the model by Eaton and Kortum (2001), suggests that there are effects from the opening of economies on innovation that goes in opposite directions. The accession into the EU inner market increases the potential profits from investing in R&D because firms can reach out to a larger group of consumers. At the same time, the new entrant firms experience more competitive pressures from other firms within the Union, and there are less room to improve upon the existing stock of knowledge in order to produce marketable ideas. Thus, there seems to be a need for an empirical analysis to determine which effects that are dominant. In that sense, the research addresses an important policy question because significant results would indicate the importance of being part of a Union on firms willingness to innovate.

I will examine this relationship using panel data. In particular, I will use Community Innovation Survey data on enterprises innovation activity, provided by Eurostat. I proceed by estimating fixed effect regressions. These empirical models are used to find difference-in-difference estimates, where I consider EU accession to be the treatment. Thus, I place firms in countries that entered the EU at one point during the available data in the treatment group, while firms in the other countries in the control group. The main model considers the effect of EU accession averaged over the years, while the alternative regression considers the effect in every period by including an interaction term. I find that EU membership results in a percentage point decrease in firms that innovate after their respective countries entered the Union compared to the mean change for the other firms. The results are somewhat robust by adding additional controls, and it shows at least a temporary negative effect of EU membership on innovation.

My research is related to other studies which sets out to identify the determinants of innovation, such as the empirical work by Aghion et al. (2005) and Acemoglu and Linn (2004). The former focuses on the degree of competition in the market, while the latter

---

1Eurostat is the statistical office of the EU. They provide high quality statistics at European level.
models market access as the main determinant. At the same time, there are other ways of modeling innovation where trade and other economic factors are important components. As I will explain, the EU membership can influence many of those factors. Thus, my work is complementary to a broader research that uses general equilibrium models to describe the interaction between trade, technology diffusion and innovation. Among those are the theoretical model derived by Eaton and Kortum (2001). It incorporates both market access through international trade and competition in a general equilibrium model to determine if research intensity does depend on opening up to trade. Other prominent models with endogenous technological change is the model by Krugman (1979) and Eaton and Kortum (1999). There are several different measures of innovation that are frequently used in economic research, see Keller (2004). Instead of using popular measures like R&D expenditure or patenting activity, I will use survey questions concerning innovation in my econometric analysis. My work contributes to the field in the sense that there are little or no previous work that sets out to analyze this research question by using several of the Community Innovation Surveys. To my knowledge, there hasn’t been any studies making use of all of these surveys to conduct a panel data analysis, despite the extensive researching on the field.

One purpose of this thesis is to use some of the insights from the research mentioned above, at the same time as I derive results from a statistical model that never have been used on the data I have in hand. I proceed as follows. Section 2 provides a background to the history of the countries that joined the EU at one point during the available data. This will help the understanding of the underlying economic mechanisms, and also justify the use of the theoretical model presented in section 3. Section 4 presents the data and descriptive statistics. Section 5 describes the statistical method used, moreover it contains the empirical results and a discussion of the validity of the study. Finally, section 6 concludes.
2 Background

To do an empirical analysis on how EU membership affects innovation, it is important to be aware of the economic impacts this membership had on the selected countries. Endogenous growth theory with models on technological change tend to incorporate many economic factors. Thus, the economic mechanisms require a comprehensive explanation. Therefore, it will be helpful to go through some of the policy changes that applied to the new member countries, and in turn evaluate the effect the policy had on certain economic variables such as foreign direct investment (FDI), trade agreements, migration etcetera. At the time of accession, firms are obliged to comply with EU law. The laws evolve around the principle of free movement of people, goods, services and capital. Compliance as in following EU directives enables the new entrant firms’ to enter the inner market with about 500 million people, and the internal policies can have an impact on firms’ position in the market.

There are many benefits associated with these policy changes. Citizens from any member country are free to work anywhere within the EU without need for visas, permits or other controls. The European Commission negotiates trade agreements on behalf of the member countries. They represent all 28 countries when it comes to negotiations through the World Trade Organization (WTO) or trade negotiations that are independent of this organization. Thus, only the EU, and not the individual member countries, can legislate on international trade deals. Moreover, intra-EU trade is characterized by free movement of goods and any quantitative restrictions on imports, exports or goods in transit are prohibited by articles 34 to 36 of the Treaty, see European Commission (2010). The inner market also benefits open economies by allowing multinational enterprises (MNEs) to locate anywhere within the Union. The European Commission emphasizes the role of a market based economy, based on principles of fair competition. They want to ensure that there is a level playing field for all firms operating in the inner market. They do so by introducing antitrust laws in order to deal with competition issues between firms in Europe. In short, these regulations prohibits the illegal creation of cartels, which may involve price-fixing or market sharing. More generally, the laws limits the market power of a overly dominant firm in order to avoid unfair pricing.\textsuperscript{2}

Preparation to enlargement took some years for the EU-10 as they had to transform their centrally planned economies into functioning market based ones in order to have a smooth transition into the European Union. They eventually succeeded in this criteria for EU

\textsuperscript{2}These rules are stated in article 101 an 102 of the Treaty.
membership, and the Commission evaluated their economies as robust in coping with the competitive pressures and forces within the Union. Whether the new member countries in practice succeeded in these criteria and managed to increase their innovation activity depends upon the memberships’ effect on their real economy.

Prior to the enlargement, the 10 countries that achieved accession in 2004 were already included in the WTO. European Commission (2006) explains that barriers to trade, FDI and other capital movements had already been largely removed. For instance, trade in goods increased rapidly in the 1990s, the decade where the European agreements were signed. This established free trade between the EU and the EU-10 countries. As a consequence, the set of all the old member countries (EU-15), became the leading trading partner for all the new countries. For example, Cyprus where included in the WTO in 1995 and by the end of 1997 they adopted the EU’s Common Customs Tariff (CCT) (World Trade Organization, 1997). The CCT is common to all member countries. Thus, the important thing to note is that the major impact of trade happened some years before the new member countries gained accession to the Union. One can argue that the possibility of experiencing more trade liberalization is present in future periods, partly if they choose to commit to adopting the common euro currency and in that way eliminate exchange rate uncertainty, lowering transaction costs and increasing price transparency. Pierce and Schott (2016) argues that the bilateral trade between the US and China experienced a boost because of the US granting of Permanent Normal Trade Relations (PNTR) to China. The PNTR removed the uncertainty associated with trade and investments with China because there was no option to increase the low tariff rates. The same uncertainty considerations are equally applicable to the European Union. Although the new member countries established trade agreements some years before accession, it was not until they became a part of the Union that the potential uncertainty associated with possible tariff increases were removed, because of the common tariff. This reduction in uncertainty could have beneficial effects on trade between old and new member countries, and potentially boost investment from established EU inner market firms because it increases the incentive to incur a fixed irreversible cost of shifting operations to new member countries. At the same time, it increases the incentive for new member countries to operate in the inner EU market, and by that increasing the competition for the already established producers.

Migration is the economic factor that got influenced the most, compared to the situation before enlargement. There is evidence of migration flows from new to old EU countries. The large differences in per capita income and wages across the enlarged Union are the major driver for this labor mobility, (European Commission, 2006). Kahanec and
Zimmermann (2010) refers to aggregate data from Eurostat, which illustrates that the new member countries experienced a drop in unemployment and increased wages in the post-enlargement period. Considering standard theory, this is not surprising because the sending countries had a negative shift in labor supply. Thus, they face reallocation in the market for labor, in particular considering the shortages of labor supply. To take one example, data from Poland documents that around 80% of the job seekers did not match labor market requirements in the manufacturing, trade and construction sectors, (Kahanec and Zimmermann, 2010). Some old member countries imposed additional restrictions in fear of mass migration from the new member countries. From a theoretical point of view this is certainly a real concern, in that a large labor supply shock can put downward pressure on wages in the old EU countries. It turned out that this concern was exaggerated in that it did not have a significant influence on labor market conditions. Statistics show that the majority of the immigrants manages to successfully integrate into the labor markets partly because of the composition of the workers, many of them taking jobs with a lower skill level compared to their education. Many of them managed to fill in the gaps in the labor markets in the receiving countries, in that they found jobs in sectors with high demand for their labor (Kahanec and Zimmermann, 2010). However, it is noted in European Commission (2006) that the numbers on migrants overestimates the actual number of new member country nationals that settles in the host country. They make the distinction in the nature of the migration flows, by noting that a large part of the workers decided to move back to their home country. This "temporary" migration distinguishes from ordinary migration in that workers flows back to domestic firms and possibly implements ideas that they have acquired abroad.

Perhaps not surprisingly, statistics show that the countries attracted more investment from abroad when they entered the Union. Bevan and Estrin (2004) found that the commission’s announcements about countries potential accession had a significant positive effect on FDI flows to those countries that where allowed to begin negotiation with the aim of joining the Union. Thus, we have reason to believe that foreign firms decide to increase their FDI into European transition economies. This increased amount of inward FDI to new membership countries is present after the public announcement, and at the time of entry. In particular, there is a clear concentration of FDI in the largest economies, that is Poland, Hungary and Czech Republic. Those countries accounted for almost 80% of the inward FDI to the new member countries in 2004. The EU-15 countries are the main contributor to this increase, with a share of 77.5% of the total inward stock in the new member countries in 2004 (European Commission, 2006). In particular, Germany was the main holder of the FDI stocks in the new member countries at the end of 2006,
owning 23% of the total EU-15 stocks (Kärkkäinen, 2008). The EU-15 firms seized the new investment opportunities by relocating parts of production to the new member countries, either by offshoring or by purchasing inputs from local producers. The magnitude of this inward FDI flows depends on the new EU countries attractiveness in term of investment. As is explained by European Commission (2006), this is partly because of the accession into the Union. However, there was still room for improvement in new member countries attractiveness in terms of investment opportunities. This is to a large extent due to the market inefficiencies and strictness of market regulations that continued to characterize the new member countries, at least relative to the other EU-15 economies. To increase the scope of the future potential FDI flows, the new member firms need to improve their competitiveness by increasing the research intensity, and create an efficient market with linkages between foreign and domestic firms.

The international mobility of capital and labor following enlargement also helped the new member countries to restructure their financial sector, in order to be more in line with the financial systems prevailing in the rest of the Union. This integration of financial markets allowed countries with undeveloped financial systems, like the CEEC countries that joined in 2004, to increase the supply of finance and expand their national financial systems. The financial integration could for example happen through foreign bank entry. The development of the banking system potentially extends the availability of credit and allow firms to lend more capital. Thus, there might be some beneficial effects from developing an efficient financial system on the country’s entrepreneurial activity (Giannetti and Ongena, 2005). This is just one possible link where accession into the EU helps promote innovation. The following section provides a formal explanation for what I believe to be the most important economic mechanisms connecting EU membership and innovation.
3 Economic mechanisms

Firms in Europe differs when it comes to the ability and willingness to innovate. The respective firms within each country are different in terms of size and profitability, and therefore have different approaches to product and process innovation. Since firms are profit maximizing, they consider profits both before and after investing in R&D. Thus, for firms to make a decision about engaging in R&D they have to consider their position in the market by investigating their competitors and be forward-looking in terms expected future profits by entering or staying in the market. Enlargement of the European Union affects certain economic variables as seen in the previous section, which in turn could induce technological chance and economic growth. Firms in new EU countries have the ability to enter new markets, and reach out to a larger group of consumers. At the same time they face increased competition from the established firms in the inner market. Improved investment opportunities in the new membership countries following their transition, makes it more lucrative to do business in these countries as indicated by the increased amount of FDI inflows. In addition, the decrease in uncertainty in terms of import tariffs helps boost trade. Thus, it is likely that the new entrants will experience increased import competition in their domestic market, and also face competition in goods markets within the Union. In my view, these are the effects that are central in the explanation as to why accession into the Union influences firms willingness to innovate.

3.1 Theoretical framework

The general equilibrium model by Eaton and Kortum (2001) considers both access to foreign markets through trade and competition with foreign inventors as factors that affects research incentives. They do so in a model that establishes the connections between forces driving innovation, and the implications of new technology for trade.\(^3\) The main workings of this model consider market access through international trade as one effect that increases innovation intensity. As I briefly explained, this is because innovation makes the firms able to produce goods at lower costs. In a trade context, this essentially means that they improve upon their comparative advantage in producing certain types of goods. Thus, they can enter markets abroad by increased export. If by entering a market with a substantial amount of demand for their product, firms might be able to expand, which in turn can increase their investments in R&D with intention of increasing productivity. Competition on the other hand, has a negative effect on research incentives because it is more difficult for firms to introduce new technology if they compete with technologically

\(^3\)The model encompasses the basic structure of other models in their earlier work. Those models were augmented in order to apply them to specific empirical research.
advanced firms that produces goods to the same market. Thus, they might choose to use existing technology and only supply to the domestic market.\footnote{Despite this intuitive explanation, Aghion et al. (2005) finds evidence that in some industries where firms are operating at similar technological levels, the competition effect is positive because firms want to innovate in order to "escape" the competition.} This model reaches an equilibrium solution for research intensity where these two effects exactly cancel. That is, increased openness of the economies, which in this case is represented as the enlargement of the EU, does not have a net positive or negative impact on research intensity. In the following, I will present the most relevant components of the model that they use to derive the result, and at the same time argue that it relates to my research question.

Countries differ in terms of their ability to efficiently produce goods. Country $i$ produces good $j$ with efficiency $z_i(j)$, where this good is among the varieties $j \in [0,1]$. The technology frontier that emerges from these efficiencies is considered as realizations of a random variable with distribution:

$$F_i = Pr[Z_i \leq z] = e^{-T_i z - \theta} \tag{1}$$

The $T_i$ is the location parameter of the distribution which indicates each country's accumulation of technology. That is, countries with high $T_i$ tend to produce any good $j$ efficiently. This is a key variable that describes the stock of ideas that have reached each country up to that point in time. The $\theta$ parameter determines the variation around the mean of the distribution. If this is low, then the variance increases. That is, the country produces some varieties with high efficiency, while others with low efficiency. This distribution determines the patterns of trade in a world characterized by perfect competition where a country $n$ buys from country $i$ only if this country is the lowest cost producer of that particular good. That is if country $i$ is the cheapest source for that good. The cost of buying from country $i$ is:

$$c_{ni}(j) = \frac{w_i d_{ni}}{z_i(j)} \tag{2}$$

Where $d_{ni} \geq 1$ represents trade costs such as tariffs or transport costs modeled in the usual iceberg cost way. Also, $w_i$ denotes the wage rate in country $i$. Since the efficiency of producing goods are modeled as a random variable, the lowest cost of importing from the cheapest source for country $n$ is also a realization of a random variable. It can be shown that it is given as:

$$G_n = 1 - e^{-\Phi_n e^\theta} \tag{3}$$

Where $\Phi_n = \sum_{i=1}^{N} T_i (w_i d_{ni})^{-\theta}$. This parameter determines how a country $n$ can make use of new technology by engaging in international trade. Eaton and Kortum uses this cost
parameter to determine the probability that country $i$ can deliver the good at the lowest price to country $n$. This probability is country $i$’s share of the cost parameter:

$$
\pi_{ni} = \frac{T_i(w_i d_{ni})^{-\theta}}{\sum_{k=1}^{N} T_k(w_k d_{nk})^{-\theta}} \tag{4}
$$

We see that it is more probable for a member country $i$ to be the cheapest source of a particular good if there are low barriers to trade between the two countries, relative to the cost of doing trade with other countries. That is, if $d_{ni}$ are close to unity. The probability could also be higher if the new member country $i$ have accumulated a large stock of knowledge relative to other countries, which means that they can produce any good with high efficiency. Lastly, it could be higher due to relative low labor costs.

The technology frontier evolves over time as an underlying process of innovation, hence $T$ can be seen as a function of time. This process involves R&D efforts from researchers that draws ideas with a Poisson rate $\alpha_i$ about how to produce a good. The ideas are realizations of two random variables. First, as to what good the idea applies. Then, the efficiency $q(j)$ with which it enables good $j$ to be produced. This is drawn from the Pareto distribution $H(q) = 1 - q^{-\theta}$. If the parameter $\theta$ is large, the probability of not improving the efficiency of producing good as a result of the new idea $j$ is large. If it is small, then there is a high probability that new ideas produces efficient technology that surpasses the current state of the art $z_i(j)$. The stock of technology or knowledge in country $i$ depends on the country specific research productivity $\alpha_i$ and on the amount of researchers $R_{it}$ in the labor force. The model does not consider technology diffusion, so to be precise, the stock of technology depends on the history of domestic research efforts:

$$
T_{it} = \alpha_i \int_0^t R_{is} ds \tag{5}
$$

An important property of the model, is that only a small fraction of the ideas are able to shift the technology frontier by surpassing existing technology. The dynamic process of innovation makes it more difficult to improve upon already existing technology that lies on the frontier. Thus, when the stock of knowledge increases, it gets harder to come up with new ideas that are better than the existing ones. Even if a firm in a new EU country manages to produce an idea that is superior to the one on the domestic market, it might not be implemented because of competition from abroad. That is, the idea is not good enough to compete with the current state of the art prevailing in the inner EU market. Eaton and Kortum considers the probability that an idea in country $i$ will be competitive in the inner EU market with a margin $m$. By using the cost expression in (3) and integrating over the idea quality distribution, it can be shown that this probability
is given by:

$$b_{nit}(m) = \int_{1}^{\infty} [1 - G_{nt}(mw_{it}d_{nt}/q)]dH(q) \approx \frac{1}{\Phi_{nt}(mw_{it}d_{nt})^\theta}$$  \quad (6)

By using equation (4) and setting $m = 1$, it can be shown that:

$$b_{nit}(1) = \frac{\pi_{nit}}{T_{it}}$$  \quad (7)

The condition states that for a firm in a new EU member country $i$ to have a marketable idea in an old member country $n$, it needs to surpass the existing technology in the domestic market. They manage this with probability $1/T_{it}$. Conditional on achieving the local state of the art, the innovation need to be of such quality that it survives the competition from the firms in the inner market, which is equivalent to them being the cheapest source of the good to which the idea applies. They do so with probability $\pi_{nit}$.

Managing to outperform competitors in the large inner market represents a difficult task for the majority of the new entrant firms. Thus, the potential benefit of increased market access through exports for the new entrants could be offset by the difficulty of becoming a technology leader in the inner market. Imagine that firms in the new EU countries experience increased competition from technologically advanced and established firms. This technology is already of such high quality that the probability for new entrants to stay in the market decreases. It could even be difficult for them to enter the EU market in the first place. Especially in the absence of knowledge spillovers from the old member firms to the new entrants. This is perhaps easiest seen when considering equation (4). The stock of knowledge in the old member countries are relatively larger or develops faster with a higher research productivity. This contributes to decreases in the probability that the cheapest source is the new member country.

As are common for other models that incorporates endogenous research, the workers are divided between producing goods and doing R&D. The equilibrium allocation of labor is determined by equality between wage and the value of doing research, $w_{it} = \alpha_i V_{it}$. To simplify the derivation of the steady state solution for research intensity, they consider a steady state where a constant share $r_i = R_{it}/L_{it}$ of the labor force engages in R&D. Then, by differentiating equation (5), the technology stock in country $i$ changes over time accordingly:

$$\dot{T}_{it} = \alpha_i r_i L_{it}$$

They show that prices are decreasing everywhere. In this model, it falls with the rate of technology. The interpretation is that prices decreases faster with higher growth of technology simply because it lowers the cost of producing the goods. Moreover, a low value for the parameter $\theta$ contributes to greater advances over existing technology, so
that changes in the price get larger. The steady state solutions are used in order to find
the value of an idea in country $i$. It can be shown that it is given as:

$$V_{it} = \frac{g_L}{\alpha_i r_i} \frac{(1 - r_i)\omega_{it}}{\theta \rho - g_L} \quad (8)$$

where $g_L$ is the constant growth rate of the labor forces everywhere and $\rho$ is the discount
rate associated with the utility function. By using that wages should equal the value of
research in equilibrium, we arrive at the expression for research intensity:

$$r_i = r = \frac{g_L}{\theta \rho} \quad (9)$$

The result shows that research intensity in country $i$ depends positively on the growth
rate of the labor force. Moreover, research intensity is greater at lower discount rates and
with smaller $\theta$. That is, if the current inventions on average are far superior to the existing
technology. "More surprising is that research intensity does not depend on country size,
research productivity, or openness. While access to foreign markets increases the potential
profits that a successful idea can earn, competition from foreign inventions makes it more
difficult to have a marketable idea in the first place", (Eaton and Kortum, 2001, p. 753-
754). According to the implications of the model, enlargement of the Union should not
create significant changes in the innovation efforts by firms in new member countries,
considering the two effects that go in opposite directions. However, the innovation efforts
by firms in new EU countries might be altered if we consider other mechanisms that go
beyond the scope of this model.

### 3.2 Technology diffusion and knowledge spillovers

The total stock of ideas in a country is not necessarily only dependent on the history of
research efforts within the country, but also on the research provided by firms in other
countries. This is true if we consider technology diffusion. The inventions made by one
firm is not only beneficial for that firm, but also for other firms adapting the new technol-
ogy. As explained by Grossman and Helpman (1991) and Romer (1990), knowledge could
be characterized as a non-rival, partially excludeable good. Non-rival because the idea
could be applied by other firms at the same time. In many cases, it is also non-excludeable
if the inventors have difficulties in being compensated for their R&D efforts when letting
other use their ideas. Because of this, many economists believe that there is a positive
eexternality associated with more innovation. Firms can adapt technology or learn from
the firms closer to the technological frontier. The estimates from Eaton and Kortum
(1999) suggests a substantial sharing of ideas. By entering the EU inner market, there is

\footnote{The utility of a consumer in country $i$ is of Cobb-Douglas form over the varieties $j \in [0, 1]$}
a possibility that firms in that particular country can make use of new technology more rapidly. They have become a part of the market, and there is more likely that there are knowledge spillovers from the high tech firms to the laggard firms.

As noted by European Commission (2006), there is an indirect positive spillover effect to the new entrant firms. One channel where technology diffuses is through FDI, where a greater volume of FDI flows makes it more likely that local new EU firms can imitate or adopt the technology from the multinational firms. Keller (2004) discusses the significance of FDI in acquiring new technology. The spillover effect could also happen through trade in intermediate goods. Imagine, one firm has no prior knowledge of the market they are about to enter. They immediately recognize the potential profits from producing some kind of new variety, but need more information about the market or knowledge on how to develop new products. They decide to engage in trade with another supplier by buying intermediate inputs and in that way discover a new way of producing the final good. This is just one example. Another possibility is that the workers who only temporarily migrated to the other member countries learned something from interacting with the foreign firms, which they then made use of when they eventually returned to their home country. Following this reasoning, the direction of the spillovers could very well be from the multinational firms to the local firms in the new membership countries. Common for the examples above, is that they consider the stock of knowledge to be related to research efforts abroad. For instance, Grossman and Helpman (1991) relates the stock of knowledge to the number of relations a firm has with businesses from abroad.

Trade and FDI is a way for firms to acquire this knowledge because it involves contacts with foreign firms who possess advanced technological knowledge. Multinational enterprises are in general known for this high degree of knowledge, and they could raise the technology transfer if they decide to invest in the new member countries. This transfer of knowledge to the new entrant firms helps them introducing new goods and services, which not necessarily have been developed internally by the firms. However, an important thing to note is that the effect on intramural R&D effort are ambiguous because the technology diffusion allows firms to simply adopt the technology, and deciding not to devote their own resources into R&D. At the same time, the firms might experience increases in productivity, which in turn makes it easier to develop their own technology. Thus, the adoption of foreign invented technology could be the decisive factor of a breakthrough in their own inventions.

---

6Keller says that the multinational firms could equally well acquire knowledge from the host country
As is discussed by Eaton and Kortum (1999), there is a distinction between an idea that has diffused and an idea that has been adopted. In order for an idea to be adopted by firms in new member countries, it needs to be superior to their own technology. Even if they manage to discover a diffused idea with quality above their own technological frontier, the usage could be restricted if the owner, who I think of as an inventor with residence in an old member country, decide to patent the idea in the new member countries. It is not in the owner’s interest - who has incurred the R&D cost - to let others use it at no charge. Therefore, the new member firms might be unable to exploit the new technology even if they experience increased openness through trade and FDI. Even if the technology is transferable with practically no cost, the remainder of the technology might be "tacit" in nature. This type of technology is hard to implement because it is difficult to fully codify it. For example, in some cases you need specific knowledge on how to adapt technology, and it is best learned by interacting directly with the inventor. This also implies that the endowment of human capital plays a role when considering how effective countries are in adopting foreign technology. A number of studies discussed in Keller (2004) suggests that there are big differences in how effective countries are at adopting foreign technology. They suggest that a certain type of skill is required. Eaton and Kortum (1996) find that acquiring knowledge as a result of inward FDI flows is increasing in the countries level of human capital. Of course, firms in new member countries are not the same in terms of endowment of human capital. Thus, the positive effect of technology diffusion on innovation is heterogeneous across the entrant firms.
4 Data

4.1 Survey data on innovation

The dataset I am using is provided by Eurostat. It is not made publicly available, so I had to apply by submitting a research proposal in order to get access. Eurostat wanted to make sure that the research question is relevant to use on their data. They are also interested in receiving the results after the research is done. Another reason for the why the data is not made publicly available is that Eurostat needs to preserve the anonymity of the enterprises that participate in the surveys. Hence, there are some guidelines to the usage of the data. I knew that it would take a long time for Eurostat to process my research proposal. Thus, I needed to make sure that the data contained enough years for relevant countries before I submitted the application form. Fortunately, it contained enough data in order for me to conduct my empirical analysis.

The Community Innovation Survey contains micro data on randomly chosen enterprises in European countries. The survey are conducted every second year from 2000 to 2014 and covers the years 1998-2014. More specifically, the CIS 3 covers the years 1998 to 2000, CIS 4 covers the years 2002 to 2004, and so on.\(^7\) I am dealing with repeated cross section data. That is, each survey reports a set of variables, where most of them indicators, that describes the enterprises innovation activity in the years the survey covers. By using all of the surveys to construct a dataset, I am able to analyze the development, or change in innovation over time by focusing on some key variables that measure innovation. The randomly chosen enterprises are asked several questions about innovation they did in that three-year span.

There are many advantages of using a subset of variables in the CIS questionnaire as a measure of innovation. One is that they are recorded on the firm level for all participating countries, hence allowing for a detailed analyze of innovation activity in each country over several time periods. In addition, it also reports the industry classification of each individual enterprise allowing me to take into consideration the fact that for some industries it is more common to use resources on developing new technology than other industries.

Every survey contains about 150 variables, although the total amount of variables deviates from one survey to the next.\(^8\) One particular shortcoming considering the data, is that many variables have large numbers of missing values. In other words, the survey

\(^7\)See the included table in the Appendix.

\(^8\)The number of variables range from 124 in the CIS 2006 to 174 in the CIS 2014.
contains questions that many firms have chosen not to answer. My focus in the analysis, will be on the indicator variables related to product and process innovation, in addition to a variable that indicates whether or not each firm engaged in intramural R&D. There are some reasons for this selection of dependent variables. First, it is the fact that most enterprises did reply to those questions, hence the full dataset only contains a small amount of missing values for those variables. Some surveys deviates in terms of questions asked, but the variables I pick in the final sample are consistently recorded in every year the survey is conducted. Perhaps more importantly, I consider those variables as the most descriptive measures of innovation out of all the available data.

After determining the outcome variables to analyze, I proceeded by collapsing the means of the relevant outcome and control variables by country, industry and year for every survey. Then I appended the datasets for each year to get the complete panel data sample. Another considerable challenge I met while I was doing data management, was to harmonize the NACE codes because the surveys deviated from one another in terms of the recording of industry classification. For some surveys the industry classifications where recorded in larger subgroups. Again, this has to do with the preservation of the anonymity of the enterprises. Thus, I had to find the corresponding subgroups for the enterprises where the NACE two digit classification where used, and change it to the larger subgroup of industries. This was an essential part of the data management because I needed to consider the industry heterogeneity in terms of innovation activity in my analysis. A similar problem occurred with the firm size variable. For some surveys it is recorded in a different fashion than for other surveys. However, this was easier to deal with because I could find the size categories in the surveys where this variable was recorded as a string, and replace them with the corresponding numbers that where already done in the other surveys. To summarize, considerable amounts of care had to be taken in the data management part of the research. The final sample is obtained by thoroughly investigate each survey, in order to keep variables that are fit to be used in a panel data analysis.

4.2 Descriptive statistics

Table 1 shows summary statistics for the variables I will use in my statistical analysis. From the original data, enterprises are asked questions regarding innovation. Since many of these questions require a yes or no answer, the original variables are indicators. Considering the product innovation variable, the enterprises are asked if they introduced a new or significantly improved good or service. Correspondingly, the process innovation variable records if the firm introduced a new or significantly improved methods of manu-
facturing or producing goods or services. In both cases, the innovations had to be new to the firm, but not necessarily to the market. In addition, the innovations could have been originally developed by the participating enterprise or by other enterprises. Considering intramural R&D, the enterprises are asked if they did creative work undertaken within the enterprise to increase the stock of knowledge for developing new and improved products and processes. I will use the mean of those variables as my outcome variables, calculated for each country-industry pair over the available time period.

Since the enterprise identifiers are removed from the data to preserve anonymity, the country-industry id is used as the panel variable. This variable considers all country-industry pairs in the dataset. The sample contains 410 pairs in total, as indicated by the value of n. Furthermore, the outcome variables are the mean of product innovation, process innovation and engagement in intramural R&D. Their respective overall means are 0.19, 0.20 and 0.36. These are the statistics after removing observations for Norway in 2008 and 2010. For these observations, all enterprises that did not do product or process innovation did not respond to the survey questions regarding the product and process variables. In other words, there were large amounts of missing values which made the means equal to 1. This could potentially have adverse effects on the estimation results.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Panel</th>
<th>Mean</th>
<th>Sd</th>
<th>Min</th>
<th>Max</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>year</td>
<td>Overall</td>
<td>2007.91</td>
<td>4.44</td>
<td>2000</td>
<td>2014</td>
<td>N = 1554</td>
</tr>
<tr>
<td></td>
<td>Between</td>
<td>2.72</td>
<td>2000</td>
<td>2014</td>
<td>n = 410</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Within</td>
<td>4.02</td>
<td>1999.11</td>
<td>2015.91</td>
<td>T = 3.79</td>
<td></td>
</tr>
<tr>
<td>product</td>
<td>Overall</td>
<td>0.19</td>
<td>0.16</td>
<td>0</td>
<td>1</td>
<td>N = 1546</td>
</tr>
<tr>
<td></td>
<td>Between</td>
<td>0.15</td>
<td>0</td>
<td>0.76</td>
<td>n = 410</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Within</td>
<td>0.08</td>
<td>-0.11</td>
<td>0.99</td>
<td>T = 3.77</td>
<td></td>
</tr>
<tr>
<td>process</td>
<td>Overall</td>
<td>0.20</td>
<td>0.14</td>
<td>0</td>
<td>1</td>
<td>N = 1546</td>
</tr>
<tr>
<td></td>
<td>Between</td>
<td>0.12</td>
<td>0.02</td>
<td>0.63</td>
<td>n = 410</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Within</td>
<td>0.08</td>
<td>-0.43</td>
<td>0.83</td>
<td>T = 3.77</td>
<td></td>
</tr>
<tr>
<td>intramural</td>
<td>Overall</td>
<td>0.36</td>
<td>0.23</td>
<td>0</td>
<td>1</td>
<td>N = 1542</td>
</tr>
<tr>
<td></td>
<td>Between</td>
<td>0.20</td>
<td>0</td>
<td>1</td>
<td>n = 410</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Within</td>
<td>0.15</td>
<td>-0.24</td>
<td>0.86</td>
<td>T = 3.76</td>
<td></td>
</tr>
<tr>
<td>size</td>
<td>Overall</td>
<td>0.55</td>
<td>0.26</td>
<td>0</td>
<td>2</td>
<td>N = 1299</td>
</tr>
<tr>
<td></td>
<td>Between</td>
<td>0.24</td>
<td>0</td>
<td>2</td>
<td>n = 385</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Within</td>
<td>0.11</td>
<td>-0.18</td>
<td>1.07</td>
<td>T = 3.37</td>
<td></td>
</tr>
<tr>
<td>cooperation</td>
<td>Overall</td>
<td>0.38</td>
<td>0.18</td>
<td>0</td>
<td>1</td>
<td>N = 1541</td>
</tr>
<tr>
<td></td>
<td>Between</td>
<td>0.17</td>
<td>0</td>
<td>1</td>
<td>n = 410</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Within</td>
<td>0.11</td>
<td>-0.07</td>
<td>0.87</td>
<td>T = 3.76</td>
<td></td>
</tr>
</tbody>
</table>

Note: The table includes only relevant variables that are used in the analysis.

---

\(^9\)Including software development in-house that meets this requirement.
especially since all enterprises from Norway belongs to the group that are not affected by the EU accession. There did not seem to be an issue with non-response enterprises for other countries.

A useful feature of the table is that it not only reports overall standard deviation, but also within and between variation. It appears that a large part of the variation in the dependent variables comes from between variation, i.e. variation between the country-industry pairs. It is made clear from the table that I am dealing with an unbalanced panel.\footnote{The Appendix contains a list of countries in the data, and which surveys they participate in.} We see this by taking a look at the year variable, and noting that there is some between variability in this variable. Moreover, the average number of years a country-industry pair was observed in the outcome data was around 3.8, as shown by the value of $T$ on the year variable. This illustrates that the panel is quite unbalanced because I have in total 7 time periods and the maximum number of years a pair was observed is actually 7.

Another important aspect is the firm size. Is it true that innovation activity is promoted by large firms, or is it highly dependent on the industry they operate in?\footnote{This was empirically tested by Acs and Audretsch (1987).} Fortunately the dataset contains firm size variables, and I choose to include that variable in the dataset and later specify it as a control variable in the regression. The variable is recorded in categories after number of employees at the end of each reference period. That is, enterprises with less than 50 employees are assigned the minimum value of 0. Enterprises that have between 50 and 250 employees are recorded to be of size 1, and lastly enterprises that have more than 250 employees are assigned the maximum value. Moreover, I intend to use the cooperation variable as an additional control variable. This variable indicates whether firms have cooperation agreements in doing R&D, and the overall mean value states that 38\% of the firms in this dataset cooperate with other firms.

The data also contains a variable for R&D expenditure as percentage of total turnover, although this is not the case for all surveys.\footnote{In CIS3, CIS4, CIS 2006 and CIS 2008 the R&D expenditure variable is recorded in national currencies.} The histogram in figure 1 provide a further illustration of the enterprises innovation effort. In particular, it shows that there is a considerable amount of heterogeneity of enterprises in terms of their ability or willingness to incur the costs of engaging in R&D. Second of all, the distribution is highly skewed, and points out that only a small fraction of the enterprises choose to allocate their resources into R&D. This is made even more clear considering the means of the selected innovation variables in table 1.
Figure 1: Histogram of R&D expenditure by enterprises in 2012

Note: The tail of the distribution is truncated for values above 0.5
In the empirical analysis, membership in the Union is considered as the treatment. Thus, I assign firms in the new member countries in the treatment group while the other firms in the control group since they are not directly affected by the policies that come along with the accession. In the following analysis, I will be concerned with differences in mean outcomes for the treatment and control group before and after the treatment is received. Therefore, I add another table that display those means over time for easy comparison between the groups.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>product</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>0.22</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.16</td>
<td>0.14</td>
<td>0.12</td>
</tr>
<tr>
<td>Control</td>
<td>0.36</td>
<td>0.33</td>
<td>0.21</td>
<td>0.22</td>
<td>0.30</td>
<td>0.18</td>
<td>0.21</td>
</tr>
<tr>
<td>process</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>0.20</td>
<td>0.15</td>
<td>0.19</td>
<td>0.21</td>
<td>0.18</td>
<td>0.14</td>
<td>0.15</td>
</tr>
<tr>
<td>Control</td>
<td>0.32</td>
<td>0.35</td>
<td>0.23</td>
<td>0.23</td>
<td>0.28</td>
<td>0.17</td>
<td>0.20</td>
</tr>
<tr>
<td>intramural</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>0.25</td>
<td>0.19</td>
<td>0.26</td>
<td>0.34</td>
<td>0.50</td>
<td>0.44</td>
<td>0.42</td>
</tr>
<tr>
<td>Control</td>
<td>0.36</td>
<td>0.34</td>
<td>0.24</td>
<td>0.34</td>
<td>0.36</td>
<td>0.44</td>
<td>0.46</td>
</tr>
<tr>
<td>size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>-</td>
<td>0.59</td>
<td>0.55</td>
<td>0.56</td>
<td>0.53</td>
<td>0.51</td>
<td>0.47</td>
</tr>
<tr>
<td>Control</td>
<td>-</td>
<td>0.61</td>
<td>0.53</td>
<td>0.52</td>
<td>0.59</td>
<td>0.55</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Note: The size variable in the first survey is recorded in a way that is not comparable with the recordings in the other surveys in terms of the values assigned to the enterprises. This is why I do not have values for this variable in 2000.

First thing to note is that a larger percentage of firms in the control group say they did do product or process innovation in every year, although this difference vary a bit over time. The same is true for intramural R&D, except for one year. I think it is important to emphasize that I am not concerned with differences in levels when doing a difference-in-difference estimation, but rather comparing the changes over time. Nevertheless, table 2 illustrates that research activity is on average higher in the countries that did not change their union status in the sample period, that is for firms in old or non member countries. I also note that firm size is about the same for the two groups in every year.
5 Empirical analysis

5.1 Methodology

The statistical method chosen is difference-in-difference (DiD) estimation, in order to estimate a causal effect of EU market access on innovation. I will do so by assigning long-term member countries in a control group whereas new member countries in the treatment group. The reason for this choice of method is that I believe there could be identified a shift in innovation activity associated with the inclusion of new countries in the EU. For this purpose, I consider the fixed effects model for $\bar{D}_{kjt}$:

$$
\bar{D}_{kjt} = \beta_{kj} + \delta_t + \eta_{EU_{jt}} + \varepsilon_{kjt}
$$

Where the mean of the error term conditional on the EU variable is assumed to be zero, which implies that they are uncorrelated:

$$
E(\varepsilon_{kjt}|EU_{jt}) = 0
$$

I will specify the dependent variable as the mean of some innovation indicator mentioned above, where $\bar{D}_{kjt}$ is the average innovation for the enterprises in country $j$, in industry $k$, at time $t$.

The dummy variable $EU_{jt}$ indicates whether or not the country entered the European Union during the time periods I have available in the dataset. That is:

$$
EU_{jt} = \begin{cases} 
1 & \text{for new member countries in the year of entry and post-treatment years} \\
0 & \text{if otherwise} 
\end{cases}
$$

I have also included time and country-industry fixed effects. The $\beta_{kj}$ captures the innovation variability that is between the country-industry pairs. Perhaps being in a particular country enhances the innovation activity because the government is subsidizing small enterprises who specializes in creating new technology or they have a long history of investing heavily in R&D. In other words, the variation in the fixed effect comes partly from omitted variables that vary across countries but not over time. Controlling this fixed effects is important because it makes the EU-effect on the dependent variables more accurate, hence removes some potential bias in the $\eta$ estimator. Moreover, the country-industry fixed effects control for the heterogeneity of industries. It is likely that I will observe different reports of innovation activity in the data dependent on the industry, that have no direct causal relationship with EU membership, but relates to features specific to the industry. Therefore, the industry fixed effects are added to remove the type of endogeneity that otherwise would have been present in the regression. In addition, time fixed effects $\delta_t$ are added to capture changes through time, in order to control for common
macroeconomic trends or shocks that influences the dependent variable.

To show that the model really represents a difference-in-differences estimation where I look at changes in the dependent variable over time for the two groups, I take first differences.

$$\Delta \bar{D}_{kjt} = \Delta \delta_t + \eta \Delta \text{EU}_{jt} + \Delta \epsilon_{kjt}$$

Then I have the difference-in-differences estimator.

$$\hat{\eta} = (\bar{D}_{kjt}^{\text{treated}} - \bar{D}_{kjt}^{\text{control}}) - (\bar{D}_{kjt}^{\text{treated}} - \bar{D}_{kjt}^{\text{control}}) = \Delta \bar{D}_{kjt}^{\text{treated}} - \Delta \bar{D}_{kjt}^{\text{control}}$$

Where $t \in [1, 2]$ denotes the pre- and post-treatment period for firms in the respective countries. The model specification has the advantage that the estimated coefficient is easily interpretable, and the study can be seen as a natural experiment where the treatment is as if randomly assigned to the enterprises.

Since there are more than one post-treatment period for the set of firms in the treatment group, it would be interesting to see if the effect persist after the year of entry or if it represents a temporary shift in innovation activity. For this purpose I estimate an alternative regression, where I include an interaction term between the $EU$ variable and year dummies

$$\bar{D}_{kjt} = \beta_{kj} + \eta_t (EU_{jt} \times \delta_t) + \epsilon_{kjt}$$

Where $EU_{jt}$ is specified as 1 for all new member countries in every year, and 0 otherwise. The dynamics of the $EU$-effect are captured by $\eta_t$, that is the vector of coefficients for the interaction between $EU_{jt}$ and $\delta_t$.

5.2 Results

Table 3 reports Ordinary Least Squares (OLS) estimates from equation (10). My preferred model specification is used in the last three columns, where time fixed effects are included by adding year dummies. The coefficient of interest is the EU coefficient.

The results suggest a statistically significant decline in innovation efforts by firms in new member EU countries once they officially becomes a part of the Union. In particular, the coefficient is $-0.0243$ in the main regression equation for product innovation, and $0.00355$ in the equivalent specification with process innovation as the dependent variable. Only the former are statistically significant.

The F-tests clearly rejects that the year dummies are jointly zero, indicating the importance of the time fixed effects in the model specification. The correction of the variation
Table 3: Initial results

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>product</td>
<td>process</td>
<td>intramural</td>
<td>product</td>
<td>process</td>
<td>intramural</td>
</tr>
<tr>
<td>EU</td>
<td>-0.0644***</td>
<td>-0.0316***</td>
<td>0.0792***</td>
<td>-0.0243*</td>
<td>0.00355</td>
<td>-0.0591*</td>
</tr>
<tr>
<td>(0.00690)</td>
<td>(0.00687)</td>
<td>(0.0153)</td>
<td>(0.0114)</td>
<td>(0.0102)</td>
<td>(0.0237)</td>
<td></td>
</tr>
<tr>
<td>Country-Industry fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time fixed effects</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>F-time dummies</td>
<td>20.86</td>
<td>29.63</td>
<td>42.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1546</td>
<td>1546</td>
<td>1542</td>
<td>1546</td>
<td>1546</td>
<td>1542</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

in the outcomes due to the fixed effects, reduces the absolute value of the coefficients of interest as we compare the results with and without the time fixed effects. Nevertheless, the coefficient with product innovation and engagement in intramural R&D remains significant.

Based on these results, the mean difference change in product innovation for the treatment group relative to the mean difference change in the control group is $-0.0243$, suggesting that firms that report that they did do product innovation in the treatment group compared to the control group have decreased in the post-treatment period. That is, 2.4 percentage points less firms did develop goods after the country they belong to entered the European Union relative to the mean change for the firms in the control group. The mean value for product innovation in the dataset is 0.19, and I would say that the percentage point decrease is relatively large in magnitude. The same interpretation applies when using the other dependent variables. That is, intramural R&D efforts decreases by 5.9 percentage points for the firms in the new EU countries after they received the treatment of Union membership, relative to the mean change in intramural R&D for the control group firms. This is also a quite large decrease considering the overall mean of 0.36.

Table 4 reports the results from estimating model (11). I have excluded the countries that entered after 2004, in order to get a clear picture of the impacts on the other countries. As mentioned above, I am interested in the full set of year $\times$ EU interactions, with 2000 as the base period. In this case, the years that follows are post-treatment periods, whereas 2004 is the year of entry. The estimation shows that the firms experience a negative innovation shock in the year of entry considering all three of the outcome variables. Interestingly, the negative effect disappears when looking at the first succeeding survey,
indicating that joining the Union only had a temporary effect on innovation activity. This is indicated by the insignificance of the $\hat{\eta}_{2006}$ for all three outcome variables. The magnitude is substantially reduced. In particular, we see a large drop in the magnitude of the effect already in the first post-treatment years. Considering product innovation the effects turns insignificant and reduces to -0.03 in 2006. It remains insignificant for the majority of the post-treatment years. The impact is the same for intramural R&D activity, which also becomes insignificant in the first post-treatment period. On the other hand, The DiD estimator turns positive and significant for process innovation in three of the post-treatment years, which supports the finding of an insignificant coefficient in the main model specification. This is because the main model considers the effect averaged over the years.

Table 4: Alternative specification - Interaction terms for year of entry and post-treatment years

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>product</td>
<td>process</td>
<td>intramural</td>
</tr>
<tr>
<td>2004 × EU</td>
<td>-0.103***</td>
<td>-0.0794**</td>
<td>-0.124***</td>
</tr>
<tr>
<td></td>
<td>(0.0312)</td>
<td>(0.0290)</td>
<td>(0.0325)</td>
</tr>
<tr>
<td>2006 × EU</td>
<td>-0.0308</td>
<td>0.0273</td>
<td>0.0368</td>
</tr>
<tr>
<td></td>
<td>(0.0238)</td>
<td>(0.0211)</td>
<td>(0.0360)</td>
</tr>
<tr>
<td>2008 × EU</td>
<td>0.0223</td>
<td>0.0930***</td>
<td>-0.0145</td>
</tr>
<tr>
<td></td>
<td>(0.0284)</td>
<td>(0.0214)</td>
<td>(0.0431)</td>
</tr>
<tr>
<td>2010 × EU</td>
<td>-0.0572*</td>
<td>0.00265</td>
<td>0.128**</td>
</tr>
<tr>
<td></td>
<td>(0.0257)</td>
<td>(0.0193)</td>
<td>(0.0452)</td>
</tr>
<tr>
<td>2012 × EU</td>
<td>0.0210</td>
<td>0.0567**</td>
<td>0.0528</td>
</tr>
<tr>
<td></td>
<td>(0.0245)</td>
<td>(0.0185)</td>
<td>(0.0454)</td>
</tr>
<tr>
<td>2014 × EU</td>
<td>-0.0122</td>
<td>0.0475*</td>
<td>0.0293</td>
</tr>
<tr>
<td></td>
<td>(0.0250)</td>
<td>(0.0189)</td>
<td>(0.0429)</td>
</tr>
<tr>
<td>Country-Industry fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1284</td>
<td>1284</td>
<td>1280</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
There are additional variables that could be used as control in the main specification (10). As mentioned in section 4, a factor that may influence innovation effort is the size of the firm. If that is the case, adding firm size as a control variable could increase the accuracy of the results. Therefore, the following regression provides a robustness check to see if the OLS assumption is reasonable and whether it is safe to do inference based on the estimates in the main regression.

$$\bar{D}_{kjt} = \beta_{kj} + \delta_t + \eta EU_{jt} + \gamma \bar{X}_{kjt} + \varepsilon_{kjt}$$  \hfill (12)

Here, $\bar{X}$ represents control variables. First of all, I am interested to do a robustness check with only one control variable. I select firms size as an appropriate control, thus $\text{size}_{kjt}$ corresponds to the control variable term in equation (12). This variable represents the average firms size, reported as the number of employees as explained in section 4.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>product</td>
<td>EU</td>
<td>product</td>
<td>intramural</td>
<td>EU</td>
<td>product</td>
<td>intramural</td>
</tr>
<tr>
<td></td>
<td>-0.0213*</td>
<td>-0.0114</td>
<td>-0.116***</td>
<td>-0.0175</td>
<td>-0.00924</td>
<td>-0.103***</td>
</tr>
<tr>
<td></td>
<td>(0.00987)</td>
<td>(0.0115)</td>
<td>(0.0243)</td>
<td>(0.0101)</td>
<td>(0.0112)</td>
<td>(0.0232)</td>
</tr>
<tr>
<td>size</td>
<td>0.0644*</td>
<td>0.102**</td>
<td>0.0300</td>
<td>0.0618</td>
<td>0.0999**</td>
<td>0.0218</td>
</tr>
<tr>
<td></td>
<td>(0.0327)</td>
<td>(0.0321)</td>
<td>(0.0564)</td>
<td>(0.0329)</td>
<td>(0.0319)</td>
<td>(0.0533)</td>
</tr>
<tr>
<td>cooperation</td>
<td>0.0650*</td>
<td>0.00397</td>
<td>0.322***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0273)</td>
<td>(0.0389)</td>
<td>(0.0593)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Regressions with additional control variables

| Observations | 1291 | 1291 | 1287 | 1286 | 1286 | 1286 |

Robust standard errors in parentheses

Note: Country-Industry fixed effects and time fixed effects are included.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In theory firm size would be an important determinant of innovation, and the results from table 3 columns 1 and 2 show that average firm size is significant. However it doesn’t alter the significance of the coefficient of interest. The difference-in-difference point estimates for product innovation is still negative with about the same magnitude as in regression (10). Even after including a control, the results still remain significant. In columns 4-6, I have included cooperation as control variable. This variable reports the average share of firms in country $j$, in industry $k$ in year $t$ that have cooperation arrangements with other firms in doing R&D. Perhaps this is an appropriate control variable. The reason being that if a firm cooperates with another firm, it is more likely that we see reports
that they did do product or process innovation. The *EU* effect becomes insignificant for product innovation, while the significance of the effect on intramural R&D remains, with an even larger coefficient than in the initial results. However, I am concerned with the fact that the original data reports a large amount of missing values for this innovation question. That is, there were many firms that did not answer this question.\(^\text{13}\) As I will discuss in the next section, I also suspect that the cooperation variable is endogeneous in relation to the outcome variables. Therefore the coefficients from the last three columns are somewhat less reliable. In spite of that, the results from only including average firm size looks promising and provides a good robustness check.

In the dataset, there are other variables that I potentially could use as controls. Unfortunately, the questionnaires differs in what questions are asked from year to year. The reason is that the surveys focus on slightly different aspects of innovation. Had the surveys been identical it would be easier to pick out potential controls to further validate my results. For example, the section concerning factors that hamper innovation such as too high costs or that the market is dominated by established firms, are not present in the surveys from 2008, 2012 and 2014. Some of those variables would have been interesting to use as controls.

To further validate my results, I am interested to see if the results are robust by specifying another treatment group. Empirical research on international trade emphasizes that bilateral trade flow is negatively associated with the countries distance to each other.\(^\text{14}\) As I argued in section 2, the new member entrants can experience increased intra-EU trade following the permanent removal of barriers to trade. At the same time, not all new countries experience the same boost in trade because of the geographic distance to the other European countries. Germany is the member country with the highest level of intra-EU trade in 2017, contributing 22.4% of the European Unions exports to other member countries, and imports 20.8% of the total imports from the EU-28 (Eurostat, 2017). The trade flows from Germany could be larger to the new member countries located closer to Germany’s economic center, while firms in countries that share a greater distance to Germany will probably not experience the same boost in trade flows. Therefore, it could make the effect of EU inner market accession on innovation heterogeneous between the treatment group firms. I want to control for this, and see if the results remain significant if I only include a subset of the treatment group. In particular, I exclude the firms from

\(^{13}\)About half of the total observations are missing for the cooperation variable. This could be a problem when considering the mean of this variable if for example many non-respondents did not cooperate, resulting in a too large mean value

\(^{14}\)A negative and significant coefficient on distance is for example found in Bergstrand (1985)
countries which economic center is outside a 800 km radius from Berlin. Thus, I remove treatment group firms from Bulgaria, Cyprus, Estonia and Romania. The results from estimating the same model as in equation (10) are displayed in table 6. Changing the treatment group have severe effects on the significance of the coefficients on the innovation outcome variables. The coefficients in column 1 and 3 are no longer significant at the 5% level as they were in table 3. This indicates that excluding control group firms from countries that lies farther from the economic center of Europe, results in an insignificant difference-indifference estimator. By comparing the tables, we see that the robust standard errors of the point estimates are noticeably larger. This is due to the decrease in the number of observations. We see that it decreases from 1546 to 1171 for product and process innovation, and from 1542 to 1168 for intramural R&D.

Table 6: Additional robustness check

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EU</td>
<td>0.0311</td>
<td>0.00671</td>
<td>-0.0343</td>
</tr>
<tr>
<td></td>
<td>(0.0172)</td>
<td>(0.0144)</td>
<td>(0.0285)</td>
</tr>
<tr>
<td>Observations</td>
<td>1171</td>
<td>1171</td>
<td>1168</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses

Fixed effects are included.

* p < 0.05, ** p < 0.01, *** p < 0.001

Finally, I am interested to check if the results from the alternative regression in equation (11) remain if I keep the changes in the treatment group. Again, I have excluded the countries that obtained membership status after 2004 in order to find the correct effects for the EU-10 countries. By looking at the results from table 7, we see that the negative effect in the year of entry is still present even though I have changed the treatment group by removing the firms that are more remote with the rest of the Union. The point estimates are still equally significant in 2004. Moreover, the EU effect turns insignificant in most of the post-treatment years, just as in table 4. The temporary negative effect remains significant and I argue that it is robust because the two different specifications of the treatment group show similar significant estimates in the year of entry.
Table 7: Robustness check with interaction terms

<table>
<thead>
<tr>
<th>Year</th>
<th>(1) product</th>
<th>(2) process</th>
<th>(3) intramural</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004 × EU</td>
<td>-0.104***</td>
<td>-0.0836**</td>
<td>-0.117***</td>
</tr>
<tr>
<td></td>
<td>(0.0314)</td>
<td>(0.0293)</td>
<td>(0.0341)</td>
</tr>
<tr>
<td>2006 × EU</td>
<td>-0.0408</td>
<td>0.0140</td>
<td>0.0272</td>
</tr>
<tr>
<td></td>
<td>(0.0241)</td>
<td>(0.0215)</td>
<td>(0.0393)</td>
</tr>
<tr>
<td>2008 × EU</td>
<td>0.0195</td>
<td>0.0656**</td>
<td>-0.00418</td>
</tr>
<tr>
<td></td>
<td>(0.0288)</td>
<td>(0.0213)</td>
<td>(0.0448)</td>
</tr>
<tr>
<td>2010 × EU</td>
<td>-0.0677**</td>
<td>-0.0183</td>
<td>0.155**</td>
</tr>
<tr>
<td></td>
<td>(0.0252)</td>
<td>(0.0196)</td>
<td>(0.0468)</td>
</tr>
<tr>
<td>2012 × EU</td>
<td>0.0262</td>
<td>0.0654***</td>
<td>0.0880</td>
</tr>
<tr>
<td></td>
<td>(0.0251)</td>
<td>(0.0183)</td>
<td>(0.0472)</td>
</tr>
<tr>
<td>2014 × EU</td>
<td>0.00184</td>
<td>0.0744***</td>
<td>0.0170</td>
</tr>
<tr>
<td></td>
<td>(0.0257)</td>
<td>(0.0186)</td>
<td>(0.0442)</td>
</tr>
</tbody>
</table>

Observations 1117 1117 1114

Robust standard errors in parentheses

Country-industry fixed effects and time fixed effects are added

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
5.3 Validity of the study

My regressions show promising results, considering the main model and the corresponding robustness checks. In any case, I need to discuss the potential threats to the validity of the study. One of the difficulties lies in the fact that market size easily can be thought of as endogenous. By modeling larger markets as something that causes a change in innovation might produce misleading results because better products tend to have larger markets (Acemoglu and Linn, 2004). In other words, innovation could cause larger markets. In this particular case, more innovation at the country level may cause a country to be a more attractive candidate for EU membership. This type of endogeneity in the treatment itself leads to simultaneous causality. This is troublesome because the OLS regression picks up both effects, resulting in a biased and inconsistent estimator. Suppose there is an omitted factor that leads to high innovation levels. It might be more likely that a country enters the EU because of this, if there actually is simultaneous causality. Hence a positive error term increases innovation, but also leads to an increased share of countries that are members of the Union. In other words, the EU variable is positively correlated with the error term in the regression. On the other hand, I can argue that the timing of this inclusion into EU is exogenous as the countries cannot anticipate exactly when it is going to happen. If there is a positive shock to innovation this could lead to EU membership, but perhaps this only becomes the reality many years after. In addition, it is not clear if the effect from innovation to EU participation is even significant as there are many other factors that plays a role in the selection of new EU countries, e.g. the management of the political system and institutions. Thus, my results can be interpreted, knowing that the causality runs from EU membership to innovation. Again, the reason why it is important to discuss absence of simultaneous causality is that it is one of properties needed, in order to ensure that the assumption \( \mathbb{E}(\varepsilon_{kjt}|EU_{jt}) = 0 \) is not violated.

There are other properties that are necessary for the classical OLS assumption not to be violated, and by that ensure the internal validity. Based on the previous reasoning for the EU variable to be exogenous, I consider omitted variables bias as a more relevant issue in my research. Although I hopefully removed the omitted variable bias by including fixed effects, there is still a possibility of biased estimates if important independent variables have been omitted. The robustness check with firm size as the only control variable showed that nothing severe happened to the coefficients of interest, which is promising. The results shows that larger firms tend to innovate more. Although it seems reasonable to use the size of the firms as control, the weakness is that firm size could be considered endogenous. What I mean by that is that R&D intensive firms could increase in size. This leads to correlation between the error term and the size variable. I showed that by also
adding cooperation as a control variable made the coefficient on EU insignificant product innovation, with a corresponding $p$-value slightly above 0.05. However, the significance of engagement in intramural R&D is still present with both controls included. Again, there is a potential issue in determining causality. A firm that innovates, could more likely have cooperation agreements with other firms in doing R&D. In addition, there is a problem with non-response enterprises concerning this variable, as mentioned above.

Considering the other robustness check in table 6, there are no significant effects. This made quite an impact in the $p$-values, all of them well above 0.05. Therefore, I cannot safely infer that there is a significant causal effect from EU membership to innovation, at least considering the effect averaged over the years. Regardless, there seem to be a short-term negative effect considering the similarities of the significant estimates in the year of entry from table 4 and 7. I think that this provides relatively strong evidence of a temporary negative shift in innovation activity for the treatment firms.

One limitation of the DiD estimation is the restrictive parallel trend assumption. This assumption is discussed in great detail by Lechner et al. (2011). Basically, the assumption states that changes in the innovation variable should be the same for both groups had there been no treatment. Thus, for making inference based on my estimates, the trends need to be similar in the pre-treatment period for both groups. Put differently, if there had been no more countries joining the EU from 2000 to 2014, i.e. no treatment received, there need to be a constant relationship in the time series for the outcome for the treatment and control group. To my knowledge, there is no formal way of testing this, but it is common to show it graphically. Ideally I would need longer time series to show this for countries entering in 2004 because the first year of data are recorded in 2000. This is unfortunate as it represents a limitation of my analysis. I will limit my discussion of this by pointing out that the parallel trend is a special DiD assumption that need not to be violated in order to do robust inference.

5.4 Discussion of the findings

At first sight, my results seem to contradict existing empirical research. Particularly if we consider EU membership purely as a way for firms to gain market access as a result of being included in the inner market. Then we expect a positive estimator because access to the inner market increases the potential profits that a successful idea can earn for the new entrant firms, thus increases their incentive to innovate. In my view, a more correct approach would be to consider EU membership as something that affects the firms in other ways than just increased market access. Domestic firms in new EU countries might
have to contract and decrease their R&D efforts simply because they realize that increased
competition from foreign EU firms makes it more profitable to stay in business by using
existing technology. In some sense, this is in line with the implications of the model
by Melitz (2003), although he does not explicitly model innovation. Increased openness
makes it more difficult for domestic firms to stay in business because of import competi-
tion. As a result, resources gets reallocated to the firms with higher productivity while the
rest has to either contract or exit the market. If we consider average firm size in table 2,
the control group are slightly larger in almost every period, perhaps indicating that these
countries have a larger share of big multinational firms. Then, this reallocation from low
productivity to high productivity export firms could very well result in contraction for
the treatment group firms. Which in turn, negatively affects their ability or willingness
to innovate. It could also be that the new entrant firms find it difficult to comply with
all the EU laws and directives, and realize that they need to adjust to these laws before
they start exploiting the market by doing R&D.

The magnitude of the change in product innovation and intramural R&D is relatively
large. Nevertheless, the robustness check with changes in the treatment group applied
to the main model suggests that the effect is insignificant. Arguably, an explanation for
this is that the positive effect on innovation by accessing foreign markets is offset by
the competition from firms in the control group. In that way, the results are to some
degree in line with the implications from the theoretical model by Eaton and Kortum
(2001), where the conclusion is that those two effects on innovation intensity cancel each
other out. However, the negative EU-effect in the year of entry remains. I motivated
this research by pointing out the need for an empirical analysis, in determining which
effects that dominate. Considering my regressions, the net effect is negative, at least in
the year of entry. Although the European Commission have introduced antitrust laws to
cope with unfair competition, the new member firms find it more difficult to discover new
ideas when they face competition from the leading firms.

I argued that technology diffusion helps the treatment group firms via knowledge spillovers
from the multinational firms. The speeding up of technology diffusion as a result of ac-
>ess to the inner market and FDI inflows from multinational firms will presumably have
a positive indirect effect on the "catching up" firms since there are knowledge spillovers.
In my view, this contributes to the explanation for why we see a larger percentage point
decrease in intramural R&D efforts compared to the other outcome variables. The reason
is that the variables considering the introduction of new products and processes also in-
clude those products and processes developed by other enterprises. While the incentives
to do in-house R&D might decrease if firms find it more profitable to copy technology from the leaders.

Although the indirect technology diffusion effect positively influences product and process innovation, multinational firms can protect their position as leaders in the market by increasing their patenting filed to the new EU countries. They will do so in order to protect their marketable ideas by limiting the hazard of imitation. As a result, the treatment group firms are not allowed to adopt the high quality patented technology that have diffused to their country. Thus, the options they have left is to adopt non-patented ideas presumably of less quality or just rely on their domestic research productivity. The latter option is costly for the firms because I have argued that it is difficult to invent something with a high degree of novelty when the country is less developed in terms of their technology frontier. Although the invention might be new to the firm, it is harder to come up with a marketable idea. As a result, they might end up using existing technology available to the firm and not devote additional resources into R&D.
6 Conclusion

The empirical model presented here investigates if there is a relationship between EU membership and innovation using a difference-in-difference estimator. The conclusion from the main results is that there is a percentage point decrease in innovation efforts by firms in new member countries compared to the change for the control group firms that are not directly affected by the treatment of EU accession. This provides an answer to an important policy question, in that the enlargement of the EU negatively affects the new member firms innovation efforts compared to the other firms. However, the additional robustness checks suggest that I should be careful when inferring a causal effect of EU membership on innovation. In particular, the results are sensitive to changes in the treatment group, in that removing firms in countries with a greater distance to the old member countries removes the negative effect. A possible reason for this is that those firms do not experience the same boost in trade flows as the firms located closer to the economic centers of the old member countries. Even so, this robustness check applied to the alternative regression does not reject the possibility of a temporary negative EU effect on innovation in the year of entry.

To understand what is driving these results, I use the theoretical model presented in Eaton and Kortum (2001). The positive effect on innovation of increased market access is exactly offset by the difficulty of finding marketable ideas when firms experience increased competition from firms that are already established in the EU inner market. My main empirical results deviates from this theory in that the latter effect possibly dominates, while the robustness checks are more in line with the implications from this model.

Surely, the research that connects innovation to openness of economies are not fully exploited. My difference-in-difference approach could possibly apply to other economic studies which sets out to investigate this connection. Such studies could for example consider the treatment as something other than EU membership. For instance a trade agreement. In addition, my research could be extended by using a more complete dataset where countries participate in more surveys, which make the panel less unbalanced.\textsuperscript{15} It could also be extended in future research that uses succeeding surveys along with new treatment group firms from candidate countries that achieves membership status. Regardless, I would say that my results are valid considering the statistical method used and the data availability.

\textsuperscript{15}The remaining CIS data are only possible to access by visiting the Safe Centre at Eurostat's premises in Luxembourg. The reason is that this particular data are only partly anonymized in terms of preserving the enterprises’ anonymity


7 Appendix

Tables

Table 8: Countries included in the data

<table>
<thead>
<tr>
<th>Country name</th>
<th>Country code</th>
<th>Year of entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iceland</td>
<td>IS</td>
<td>-</td>
</tr>
<tr>
<td>Norway</td>
<td>NO</td>
<td>-</td>
</tr>
<tr>
<td>Belgium</td>
<td>BE</td>
<td>1958</td>
</tr>
<tr>
<td>Germany</td>
<td>DE</td>
<td>1958</td>
</tr>
<tr>
<td>Ireland</td>
<td>IE</td>
<td>1973</td>
</tr>
<tr>
<td>Greece</td>
<td>EL</td>
<td>1981</td>
</tr>
<tr>
<td>Portugal</td>
<td>PT</td>
<td>1986</td>
</tr>
<tr>
<td>Spain</td>
<td>ES</td>
<td>1986</td>
</tr>
<tr>
<td>Cyprus</td>
<td>CY</td>
<td>2004</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>CZ</td>
<td>2004</td>
</tr>
<tr>
<td>Estonia</td>
<td>EE</td>
<td>2004</td>
</tr>
<tr>
<td>Hungary</td>
<td>HU</td>
<td>2004</td>
</tr>
<tr>
<td>Lithuania</td>
<td>LT</td>
<td>2004</td>
</tr>
<tr>
<td>Latvia</td>
<td>LV</td>
<td>2004</td>
</tr>
<tr>
<td>Slovakia</td>
<td>SK</td>
<td>2004</td>
</tr>
<tr>
<td>Slovenia</td>
<td>SI</td>
<td>2004</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>BG</td>
<td>2007</td>
</tr>
<tr>
<td>Romania</td>
<td>RO</td>
<td>2007</td>
</tr>
<tr>
<td>Croatia</td>
<td>HR</td>
<td>2013</td>
</tr>
</tbody>
</table>

*Note:* Sorted after year of entry into the EU.
Table 9: Community Innovation Surveys

<table>
<thead>
<tr>
<th>Survey</th>
<th>Years covered</th>
<th>Non-participating countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIS3</td>
<td>1998-2000</td>
<td>CY, HR, IE, SI</td>
</tr>
<tr>
<td>CIS4</td>
<td>2002-2004</td>
<td>BE, CY, HR, IE, IS</td>
</tr>
<tr>
<td>CIS 2006</td>
<td>2004-2006</td>
<td>BE, DE, HR, IE, IS</td>
</tr>
<tr>
<td>CIS 2008</td>
<td>2006-2008</td>
<td>BE, EL, HR, IS, NO*</td>
</tr>
<tr>
<td>CIS 2010</td>
<td>2008-2010</td>
<td>BE, EL, IE, IS, LV, NO*</td>
</tr>
<tr>
<td>CIS 2012</td>
<td>2010-2012</td>
<td>BE, EL, IE, IS, LV</td>
</tr>
<tr>
<td>CIS 2014</td>
<td>2012-2014</td>
<td>BE, IE, IS, SI</td>
</tr>
</tbody>
</table>

Note: Countries from table 8 not listed here are participating

*Dropped because of high numbers of missing observations.