U.PORTO

FEP FACULDADE DE ECONOMIA UNIVERSIDADE DO PORTO

EXPORTER FIRMS BEHAVIOUR, EVIDENCE FROM PORTUGUESE FIRMS USING MICRODATA

Luís Pedro Manso Machado

Dissertation

Master in Economics

Supervised by

Paulo de Freitas Guimarães

2018

Biography

Luís Pedro Machado, was born on February 18, 1994, in Santa Marinha, Vila Nova de Gaia. He started the Bachelor program in Economics in 2013 at the Faculty of Economics of the University of Oporto, and concluded his degree in July 2016. In September 2016 he joined the Master in Economics in the same institution.

During the course of his academic life, it is relevant to distinguish some high points. The project done in partnership with OBEGEF, was the first experience of autonomous investigation, under the guidance of Doctor Nuno Ricardo Moreira. Also noteworthy, is the collaboration with FEP Junior Consulting, in the Finance Department. During the Master Degree, the promotion to Chief Financial Officer of FEP Junior Consulting, was one the most demanding moments, giving him the responsibility of managing a full department and all the financial aspects of the company. Also, during this year, two other achievements are worth mentioning. Firstly, the nomination as best student in the Master in Economics, in April 2017. Secondly, the first prize in the essay competition "Economia e o Futuro", giving way to his first published article, in "Cadernos de Economia", the official publication of the "Ordem dos Economistas".

Regarding professional experience, he did internships in accounting and data analytics during the first and second year of the Bachelor's and in the first year of the Master's, respectively. In September 2017, he initiated his professional life in PricewaterhouseCoopers (PwC), collaborating in the Transfer Pricing team, as Assistant Consultant.

In July 2018, he presented the Master's Thesis "Exporter firms behaviour, evidence from Portuguese firms using microdata" for conclusion of the program and acknowledgment as Master.

Acknowledgments

To all my family, friends and everyone that has impacted my life during this long period of learning and evolution. I must thank the Faculty and all its elements, from Professors to Assistants, for all the knowledge and challenges that during my course forced me to evolve and improve. Without the increasing difficulty, volume and detail of several lessons, my curiosity wouldn't have propelled me so far.

To my supervisor, Professor Paulo Guimarães, without whom this work wouldn't be possible. Thank you for all the support and help given during the elaboration of this Thesis, and for all the time spent giving feedback, and bringing new ideas and suggestions for this project.

A special acknowledgment is due to the Bank of Portugal, for the creation, maintenance and management of the database "Central de Balanços" and for the permission to base this dissertation in this database. Also, a special appreciation is due to the entire BPlim team for the assistance in all phases of the dissertation.

Lastly, I would like to thank my family and closest friends for supporting me in the harshest moments, giving me the strength to continue this challenge and completing it in a timely manner. To Rita, for all the nights badly spent, for all the periods I was working and could not be with, thank you for all your support and making sure I did not deviate from completing this project.

Abstract

The combination of a high quality and universal firm level database for Portugal allows the detailed study of firm's behaviour. We use BPlim's harmonized Central Balance Sheet panel for the period of 2006 to 2015 to evaluate the different behaviour of exporters and non-exporters in Portugal. We follow on the self-selection and learning-by-exporting literature, estimating several exporter productivity premiums. After finding solid evidence of a productivity advantage of exporters compared to non-exporters, which seems to emerge several years before firms start to export, we expand our study in order to explore the causality of the previous findings. Thus, we estimate a logit fixed effects model to assess the impact of several variables in the export propensity of a firm. We corroborate the self-selection theory, given the significance of labour productivity in probability of a firm exporting, as well, as significant effects of firm absolute size, relative market share, sector concentration and investment.

JEL Codes: C23 D22 F10 F14

Keywords: Fixed Effects Models; Panel Data; Firm Behaviour; Exports; International Trade; Self-selection.

Sumário

A combinação de uma base de dados universal e de elevada qualidade para Portugal permite um estudo detalhado do comportamento da firma. Utilizamos o painel harmonizado da Central de Balanços do BPlim para o período de 2006 a 2015 para avaliar o comportamento diferenciado entre exportadores e não exportadores em Portugal. Seguimos a literatura de *self-selection* e de *learning-by-exporting*, estimando vários *premiums* de produtividade para empresas exportadoras. Após encontrar evidência sólida de uma vantagem produtiva dos exportadores face a não-exportadores, a qual aparenta emergir anos antes das firmas começarem a exportar, expandimos o nosso estudo de forma a identificar a casualidade dos resultados anteriores. Desta forma, estimamos um modelo *logit* de efeitos fixos de forma a averiguar o impacto de várias variáveis na propensão à exportação de uma empresa. Corroboramos a teoria de *self-selection*, dada a significância da produtividade do trabalho na probabilidade de uma empresa exportar, assim como, o efeito significativo da dimensão absoluta da empresa, quota de mercado relativa, concentração do setor e investimento.

Table of contents

Table of figures

Table of tables

Chapter 1 – Introduction

The period of the Great Recession was characterized by generalized economic contraction, low economic growth rates, diminishing investment rates, contraction of private and public consumption, and an increase in public debt that constrained governmental stimuli to the economies.[1] This drop in present output is seen in classical macroeconomics as temporary, and regarded as part of the business cycle. Recent authors have found the existence of persistent long term effects on potential output, a phenomenon coined "hysteresis effects" (Ball (2014)). In a setup of lower potential growth and constrained fiscal stance, external demand can be a way to fuel economic growth and recover an economy. In a setting of reduced private consumption and lower output, imports are expected to fall, leaving only exports as the way to improve the current accounts of a country (Burda and Wyplosz (2013)).[2]

The prevalent idea in the literature is that productivity defines if firms can become exporters or not. The advent of higher quality micro level databases that provide wider coverage for individual firms or establishments made clear that there existed a high degree of heterogeneity in firms, in terms of productivity, size, as well as other measures (Mayer and Ottaviano (2008)). Trade dynamics and determinants theory has had several contributions, from Classical texts such as Adam Smith's absolute advantage and David Ricardo's comparative advantage to Heckscher-Ohlin neoclassical model. Afterwards, more sophisticated models eclipsed the previously stated ones with the capability of incorporating empirical regularities that have been detected with the use of microdata (Krugman (1980); Melitz (2003); Melitz and Redding (2014)). These new models have provided a theoretical backbone for the understanding of said empirical regularities, creating a framework which has been extended in several contributions that advance different reasons for the difference in exporters and non-exporters productivity, the effects of entry and exit of firms in the export market and the effects of trading on firms' mark-up. These models base the trade

[1] For a higher detail on the effects and causes of the 2008 financial crisis see "OECD Insights: From crisis to recovery".

[2] As seen from the base formulation of the IS-LM model in open economy, where Y = C + G + I + (X-Q), where the domestic product is calculated as the sum of private and public consumption, investment, and external demand.

determinants on different characteristics, which vary from monopolistic competition and scale economies, to heterogeneity in the firms' exogenous productivity parameter, giving indication to possible pointers that lead firms to export. Several empirical aspects of firm data and desegregated trade statistics are incorporated in the models, and explanations to the differentiated performance are advanced. These models also give rise to a new debate on the literature, that of the self-selection versus the learning-by-exporting theory (Wagner (2007)). The self-selection theory postulates that only the most productive firms are selected to become exporters (*ex-ante*), given that only the most efficient firms can surpass the fixed costs of exporting. The learning-by-exporting theory advances that firms start exporting, for example, due to an exogenous shock, and that in the process of exporting, they become more productive from the cumulative knowledge incorporated in goods traded and also in the relationships with external firms.

According to Mayer and Ottaviano (2008), a lack of statistical information on firm-level behaviour, has excluded firm-level study from the policymakers' array of analysis. Nonetheless, the authors also refer that with the emergence of high quality micro-level databases, this information should start to be included in a policymaker's analysis in order to enhance firm behaviour observation. As stressed by Oliveira (2016), Portugal has available excellent quality databases, due to their detailed input and the rigorous and long process of quality control the data is subject to. One of these databases is the "Central de Balanços do Banco de Portugal" (Central Balance Sheet Database of Bank of Portugal) where firm and establishment financial and operational level information can be obtained from 2006 to 2015, giving basis to potentially high value studies about Portuguese firms' behaviour. On the other hand, the author stresses that while this database, in specific, and others in general, are available, there are few studies using them, defeating the purpose of their creation.

The objective of this thesis is twofold: First, to assess if exporters are more productive than non-exporters; and secondly, to identify relevant variables that have an impact on the decision of a firm to export. The conjugation of these two objectives may allow for new insights relating to the behaviour of Portuguese exporters.

The thesis is organized as follows. Section 2 presents a literature review on trade participation determinants, internationalization and international trade, as well as their historical evolution, giving a quick overview on the main econometric tools to analyse these effects on real data. Section 3 presents an overview on the data from the Central Balance

Sheet Database, and uses it to provide a general description of firm organization and characteristics in Portugal. Section 4 presents, provides a brief characterization of the exporters. Section 5 presents the methodology and model selection. Finally, Section 6 presents the econometric results and their discussion, and Section 7 presents and summarizes the main findings of this investigation.

Chapter 2 – Literature review – what leads firms to export?

In this section, we review the literature on international economics and international trade. The objective of this analysis is to assess the most relevant variables affecting the change of a firm's behaviour to become an exporter, as well as, to identify the main theories for the existence of exporters.

The internationalization process of a firm is dependent on several variables, being influenced by domestic and foreign conditions, as well as by internal and external characteristics of the firms. This process does not involve only international trade flows, but also capital flows via direct foreign investment. The analysis should not be focussed only on outward flows, given that inward flows are a possible way of acquiring technology and knowledge that allows firms to expand outwards (Vagos (2015)).

The study of why countries and firms partake in economic trade, has been researched since the classical theories of Economics. Vagos (2015) and Sousa (2016) provide a historic review of the main theories defining this branch of International Economics. A beginning point on the literature starts with the seminal works of Adam Smith on absolute advantage, which advocated that countries would trade goods for which they had absolute advantage in cost or productivity, and Ricardo's theory of comparative advantage, advocating that countries would specialize and trade goods for which their opportunity cost was lower. These classical models were latter obfuscated by another seminal work, the Heckscher-Ohlin model, where international trade is defined by the factor quantities in countries and their relative price, creating the neoclassical trade model.

While providing a wealth of determinants to trade dynamics, the previous models could not justify or predict the intra industry trade pattern, in a scenario of high international trade and countries with similar factor endowments. Krugman (1980) presented a model, based on the existence of scale economies, product differentiation and imperfect competition, which evidences how and why firms in a country may export products for which there exists home market for. The framework is further improved in Melitz (2003), where export participation and performance is based on firms' productivity after an initial investment to cover a sunk cost of entering on the international market, making endogenous the entry and exit dynamics in the export market. As such, the paper gives rise to a theoretical background to a selection mechanism in the economy, where only the most efficient firms

can participate in international trade. In this model, firms face a sunk cost of entering in the export markets, only knowing after this investment if they have the needed productivity to participate in international trade, this is, there is a cut-off productivity level below which, firms exit the export market without exporting, or otherwise export until a negative shock reduces the firm productivity below said cut-off. An interesting conclusion of this model, other than the entry and exit dynamics of firms, is that these new entrants have lower productivity levels than incumbents and long term exporters, since lower productivity firms have exited the market, leaving only more productive firms. This fact is anticipated by firms, giving basis to the theory of self-selection, indicating a process by which only the most efficient firms select themselves into the export market (via investments to counteract the existence of sunk costs). A further extension of this model is presented in Melitz and Redding (2014) where there is a combination of both heterogeneous firms and product differentiation, giving a framework to analyse intra industry trade. This framework gives support to empirical findings of higher average productivity for exporters, due to firm dynamics and scale economies, given their presence in more and bigger markets.

In addition to the self-selection theory the literature has been debating the learning-by-exporting theory, which states that firms productivity increases with trade time, and as such, that firms learn from this process (Wagner (2007)). While these visions may collide, they are not mutually exclusive, as it is possible that after an effect of selection, firms continue to learn from the continuous process of exporting, leading to increases in productivity.

Bernard and Jensen (1997) are sometimes identified as the motivators of the self-selection theory, given the authors' investigation on assessing which firms start to export, and if these firms outperform non-exporters. This line of thought is supported by the idea that only more efficient and larger firms can support the added costs of doing business with agents in other countries, namely, due to transport costs, demand uncertainty and different tastes and preferences. The authors assess this behaviour by, firstly, comparing *ex-ante* productivity levels for firms, comparing future exporters and future non-exporters. This analysis leads to the conclusion that future exporters show larger levels of employment, labour productivity and pay larger wages before entering the export market (*ex-ante*), meaning that these firms have structural advantages or favourable characteristics before starting to export.

To access the causality of the phenomenon, the authors estimate a participation model, concluding that variables such as total employment and wages payed increased the probability of a firm exporting, as did previous exporting experience. On the other hand, testing for increased growth rates on firms that start export (e.g. the learning-by-exporting theory), the authors do not find conclusive evidence of this phenomenon. The authors in subsequent studies conclude for the importance of *ex-ante* characteristics over the positive effects of learning, consolidating this view.[3]

Following this seminal article and its extensions, many microeconometric studies focussed on the relationship between exports and productivity. Wagner (2007) presents a review of several articles, and concludes: 1) that exporters are bigger than non-exporters; 2) for the evidence in favour of the self-selection theory; and 3) for the lack of evidence of exporting effects on productivity. Even if these findings are robust to the use of different methodologies, the authors advance that these conclusions should not be considered stylized facts.

In order to address some of the difficulties pointed by the author, The International Study Group on Exports and Productivity was formed, producing studies based on similar methodology for 14 countries.[4] The International Study Group on Exports and Productivity (2008), addressed several key questions and presented their findings. They concluded for the existence of an exporter productivity premium, finding evidence of heterogeneity across countries. Secondly, they computed an *ex-ante* export premium to assess if future exporters present higher *ex-ante* productivity than future non-exporters. They found favourable evidence of this matter for developing countries, corroborating the self-selection theory for these countries. Lastly, *ex-post* export premium was tested, meaning a test on the positive effect of exporting on productivity. Given the lack of significance, the results were inconclusive. These findings maintain overall robustness to different specifications of productivity.

[3] See (Bernard and Jensen (2001); 2004).

[4] The group consists of teams working with comparable micro level panel data for 14 countries and a selected set of specified models. The countries analysed are the following: Austria, Belgium, Chile, China, Colombia, Denmark, France, Germany, Italy, Republic of Ireland, Slovenia, Spain, Sweden and United Kingdom.

Silva, Afonso, and Africano (2012), present an overview on the theory of learning-by-exporting as well as possible sources of knowledge spillover and subsequent incorporation by firms. External trade partners not only have specific knowledge about other markets and possibly other production techniques but also have an interest in buying products of higher quality and/or lower prices. These preferences force firms engaged in international trade to either specialize in higher quality production or to be more cost efficient, in order to respond to external demand. This effect can be accelerated by way of acquiring intellectual property (e.g. product design or patents) which gives firms new knowledge and improves their productivity. Moving to a more theoretical background, as opposed to the case study approach detailed above, the authors present the ideas of Grossman and Helpman[5], where intangible ideas spillover from the trade of tangible goods, leading to the acquisition of new knowledge by buyers. Learning-by-exporting is modelled in that theoretical model by the number of contacts between domestic firms and external ones, which is assumed to be correlated with the commercial weight for said domestic firm. With the advent of microdata the literature continued to study learning-by-exporting, reaching conflicting results. The authors also note that the estimations may capture effects of other variables, such as scale economies or learning-by-doing gains. Estimation obstacles on the use of case studies and microdata panels are discussed, with particular focus on non-observable characteristics, such as, management decisions. Also, the authors propose that Total Factor Productivity (TFP) should be chosen as opposed to simple labour productivity, which presents increased difficulty to the process, since the choice of a production model and possible endogeneity issues arise. Other presented methods involved testing the stochastic dominance of the productivity distributions of exporters and non-exporters[6] and quantile regression[7] to access the effects of different variables on different segments of the population. While different methodologies have been employed, a common problem arises, being this the existence of selection bias, given the possibility that becoming an exporter can be a non-random effect, but rather, a deliberate result of its management action.

Martins and Yang (2009), do a meta-analysis of 30 different studies in order to account for methodological differences in the estimation of possible learning-by-exporting

[5] See Grossman and Helpman (1991).

[6] See Delgado et al. (2002), Girma et al. (2003) or Cassiman and Golovko (2007).

[7] See Yasar et al. (2003).

effects on productivity. The authors detect a positive effect of exporting on productivity, being this effect larger in developing countries, leading to the understanding of the importance of openness to trade for these countries, as a way to reach the technological frontier faster. Continuing on this line, Timoshenko (2015) explores the persistence of exporters in the market, given their resilience to productivity shocks and exchange rate fluctuations, measured by the high percentage of exporters that continue to export in the following period. The author advances two possible reasons for persistence, first the existence of sunk costs to exporting, leading firms to continuing exporting since this cost will be incurred again if the firm exits and wants to enter the market again, and second, the existence of learning effects, which make older firms more profitable, leading to their persistence. The author proposes a model where learning is modelled as the age dependence of sales, measured by the impact of the number of years exporting, on export sales. He reached results that were favourable to the learning hypothesis, while showing the appearance of diminishing returns to export experience and the learning process, such as advanced by Silva et al. (2012). Also, (Timoshenko, 2015) test the effect of learning effects on export persistence and find that the marginal effects are positive and significant, and also, that the effects associated to sunk costs are considerably lower than most estimates in the literature, further consolidating the effect of learning in firms' behaviour.

The relevance of this literature from the policymakers' point of view is impactful. If more efficient firms start to export, there is no reason to support possible "future winners", as today's most efficient firms should be the ones supported (Bernard and Jensen (1997)). In the other hand, if learning effects are present, the assistance to new exporters can be determinant in assuring that future profitable firms do not exit the market prematurely (Timoshenko (2015)). These questions are increasingly relevant when empirical facts such as the ones in Mayer and Ottaviano (2008) are presented, that exports are concentrated in a few number of "superstar" firms. These are large, productive, and more capital intensive firms, opening uncertainty as to the aggregate benefit of policy. While these authors discuss some policy implications, they caution that these should be tested and observed on a case by case basis, as economies differ from one another.

Another line of research, dwells in the question of the *ex-ante* effect of participation, that is, if the Self-Selection of firms is an exogenous effect or an endogenous one. That is, if firms start exporting due to positive productive shocks or positive shocks on trade barriers

costs, or if firms intentionally try to increase their productivity before starting to export (Hallward-Driemeier, Iarossi, and Sokoloff (2002); Impullitti , Irarrazabal, and Opromolla (2012)),. This line of thought is followed by exploring if becoming an exporter is an exogenous effect (for example, due to the lowering of trade costs or due to some demand shock) that leads a firm to start exporting, or if this state change is done by a concerted *ex-ante* decision to improve productivity in order to participate in the external market. From a database of microdata from five East Asian countries, Hallward-Driemeier et al. (2002) advance that firms aiming to export make strategic decisions in order to increase their productivity, hence beginning a conscious self-selection behaviour, from investment to training decisions. The authors analyse this theme considering that the market orientation (either internal or external) in the first year of activity determines if the firm was incorporated with a vision to compete in one or the other market, as such, they assume this decision to be an exogenous one. This assumption means that if a firm starts exporting in the first year of activity, it was created with this purpose in mind. Based on this assumption, the authors identify that this conscious self-selecting firms have higher productivity levels than firms that started to export or those that never exported, indicating a possible avenue from which the authors conclude for a conscious self-selection of firms. The authors also note that firms with foreign capital ownership, and especially with higher percentages, tend to be more productive and sell more in the foreign market.

López (2009) proposes that in order to enter in international trade, firms have to present either higher quality products (product differentiation) or lower prices (more productive producers), which must be done either by investing in new equipment or in new technologies. The paper proposes a model where firms produce two products, one low quality homogeneous product, and a high quality differentiated one. The model calibrated to Chilean firm level data controlling for the internal sales growth of the sample firms after investment, shows that the indicator remains almost constant while firms start selling to the foreign market. This behaviour gives rise to evidence of a conscious goal of investing in order to export, since this new capacity is not applied to domestic sales. This paper gives evidence of a conscious self-selection (firms actively invest in order to increase productivity above the necessary threshold to enter the foreign market). Moreover, Alvarez and López (2008) had already analysed the impacts of heterogeneity and other variables in the export probability, having found that capital intensity and investment in general are key metrics that influence the probability of exporting.

Another possible determinant of exporting, can be pressure due to internal market contraction, creating the incentive for the firm to start exporting in order to maintain profitability, market share or simply its activity level. Muñoz-Sepúlveda (2014), explores this avenue of "residual sales", where firms start exporting excess production, trying to gain competitiveness and long-run success. The author also focuses on the impact in domestic sales of being an exporter, reaching a conclusion for the increase of the level of domestic sales when a firm is an exporter, while growth rates of domestic sales tend to decrease. This behaviour indicates a possible substitution effect between domestic and foreign sales. Also, evidence is shown that firms that never export tend to be more prone to internal demand fluctuations, and suffer higher growth rate decreases. Blum, Claro, and Horstmann (2013) give insight for this firm behaviour. As internal demand contracts, firms can use more of their fixed capacity to satisfy new external clients. Nonetheless, some firms leave exporting activities as the internal demand recovers, leading to a constant entry and exit from the export market by, namely occasional exporters, as opposed to permanent ("perennial") exporters. Lee, Beamish, Lee, and Park (2009) find evidence of increasing export intensity due to internal market contraction in South Korea, corroborating the findings of previous authors.

Mayer and Ottaviano (2008) advance some stylized facts about international firms based on microdata insight, such as the decomposition of aggregate international flows in two margins, the intensive margin or average flow and the extensive margin or the number of entities involved in such operation. The authors advocate for the predominance of the effect of the extensive margin, its highly skewed distribution and the characteristics of firms operating in the international market. The author show that there is a high concentration in the international market from a firm point of view, existing a dominance of external commerce by a handful of big and productive firms. This internationalized firms present higher productivity than domestic firms. This characteristics increase in magnitude when instead of considering only the relationship between exporters and non-exporters, we consider entities that perform foreign direct investment (FDI). The inclusion of FDI presents evidence that firms that operate in the external market have more favourable characteristics than domestic ones. This capacity constraint can be due to the presence of increasing short run marginal costs, as evidenced by JaeBin and Alexander (2013).

The literature tends to favour the self-selection theory over learning-by-exporting, but mixed results have been discovered, especially when resorting to individual firm data. The debate goes further than an academic argument as previously stated. If the self-selection theory is correct, policymakers should help in improving firms' productivity and/or reducing trade barriers and associated costs, in order to allow a higher number of firms to enter the market. On the other hand, if more than an exogenous shock, this productivity gain is a conscious one, then policymakers should help firms invest in order to improve their productivity or allow for conditions to improve technologies spillovers. Lastly, if learning-by-exporting is correct, policymakers should seek to assist firms in continuing to participate in external trade, trying to make firms more resilient to possible negative shocks that may cause them to participate in the market and exit (Impullitti et al. (2012)).

Chapter 3 – Data and firm general characterization

3.1. Database description

Given that the objective of this study is the identification of productivity differences between exporters and non-exporters, and also, the identification of relevant variables impacting on the decision of starting to export, firm level data has been chosen for the analysis. This decision was taken due the large amount of detailed firm level data existent, which allows for insights to be extracted on individual firm's behaviour. This chapter identifies the database used and explores some of its variables in order to present their behaviour and conclude for their possible use as exporting determinants.

The database under use is the "2017 *BPlim's Central Balance Sheet Harmonized Panel*", which is a firm level database provided by the Portuguese Central Bank (Banco de Portugal (2017)). This database is a smaller version of the institution's main *Central Balance Sheet Database*, because it includes only variables that are comparable over time, including some variables that have been harmonized to account for the change in accounting standards that occurred in 2009. [8] The database has annual information at the firm level, from 2006 until 2015, related to non-financial corporations operating in Portugal. The information in the database is self-reported by each individual firm, and is based on the mandatory Annual Declaration ("*Informação Empresarial Simplificada*" - IES) which encompasses economic and financial information for the entity, and is submitted to the Portuguese Tax Authority. Afterwards, the quality of the data is controlled by a rigorous quality control process by the Banco de Portugal.

The panel used spans from 2006 to 2015 and practically has universal coverage of all non-financial corporations operating in Portugal.[9] Non-financial corporation classification is based on the "European System of National and Regional Accounts" (ESA 2010).[10] In terms

[8] The harmonization has been done in terms of the accounting rules applicable during the period, from POC ("*Plano Oficial de Contas*") to SNC ("*Sistema de Normalização Contabilística*"), guaranteeing the comparability of the firm's information for the whole period existent in the database.

[9] The database does not cover sole proprietors. Also, in a given year companies may not be present in the database if they are late filing IES, if their main activity is finance (sector K economic classification) or if they fail to pass the data quality control implemented by Banco de Portugal.

[10] For more information, see Banco de Portugal (2017) or http://ec.europa.eu/eurostat/web/esa-2010

of sectorial information, the CAE rev.3 classification applies, representing 17 sectors of activity linked to the production of non-financial goods or services.[11] For further details on the consistency and quality control of the Central Balance Sheet Database, the attention of the reader is directed to Banco de Portugal (2017) and Oliveira (2016).

In the specific case, the dataset comprises 3,709,683 data points[12], represented by 613,107 unique firms covering the 2006 to 2015 period (10 years).[13] The average number of observations per firm is 6.05.

Table 1 - Distribution of number of observations per firm

Obs per individual	Freq	Perc.	Cum.
1	66,426	10.83%	10.83%
2	67,458	11.00%	21.84%
3	62,470	10.19%	32.03%
4	48,748	7.95%	39.98%
5	42,827	6.99%	46.96%
6	37,144	6.06%	53.02%
7	32,841	5.36%	58.38%
8	31,124	5.08%	63.45%
9	30,629	5.00%	68.45%
10	193,440	31.55%	100.00%
Total	**613,107**	**100.00%**	**-**

Table 2 –Top 10 observations patterns in the database

Pattern	Freq.	Percent	Cum.
1111111111	193,440	31.55%	31.55%
0000000001	30,953	5.05%	36.60%
0000000011	26,854	4.38%	40.98%
0000000111	26,051	4.25%	45.23%
0000011111	19,767	3.22%	48.45%
1110000000	19,468	3.18%	51.63%

[11] CAE – *Classificação de Atividades Económicas – revisão 3* (Economical activities classification – revision 3)

[12] User defined variables were created in order to specify, for example, variable growth rates or financial ratios. These are presented in Appendix 1.

[13] The database was cleaned of "ghost" firms. This "classification" was based on an *ad-hoc* filter which excluded firms with less than a thousand euros in sales or assets and with zero employees ("ghost" firms). Also one firm with constitution year of 2017 was eliminated since this was considered faulty reporting. This process eliminated 957 firms, which represented 1,311 observations and less than 0.1% of the observations of the database.

Pattern	Freq.	Percent	Cum.
0000001111	19,451	3.17%	54.80%
1100000000	19,431	3.17%	57.97%
1111000000	17,134	2.79%	60.76%
1000000000	16,521	2.69%	63.46%

From the preceding tables it is possible to see that, only 32% of firms are associated with a low number of observations (3 or less observations), while a significant part of firms have information for the full period – 31.6% of the individuals. Also, from the analysis of Table 2 it is possible to conclude that the most common patterns are for firms observed for continuous periods. These top patterns lead to the conclusion that the majority of the firms are active for a continuous period, being of reduced weight those that either enter and exit from the active market, or that show activity only in the middle of the period under consideration.

3.2. Observations dynamics

Table 3 - Flows in number of firms in database

Period	Total	First	Last	Sing
2006	343,623	343,623	16,521	16,521
2007	358,499	34,074	22,317	2,886
2008	368,367	32,772	25,862	2,746
2009	368,119	26,999	25,640	2,165
2010	368,549	26,547	23,352	1,945
2011	377,014	30,458	28,039	2,500
2012	376,842	26,897	27,783	2,072
2013	380,899	31,385	27,525	2,093
2014	383,093	29,399	31,390	2,545
2015	384,678	30,953	384,678	30,953

Table 3 decomposes the flows in the firm identifier variable in different categories, as follows: 1) *Total*, indicates the total number of distinct firms in each year; 2) *First*, indicates the number of firms which were observed for the first time in the period. In the first period of the database, all observations were observed for the first time; 3) *Last*, identifies the number of firms which were observed for the last time in a time period; and 4) *Sing*(*leton*), shows the number of firms which are only observed at one time period, meaning, firms that entered and exit in the same year. Firms may stop being observed mainly because of three motives. First, they may cease their economic activity, second, their reporting may be considered of low quality, thus being excluded by the quality control procedure of the Banco

de Portugal, and third, their main activity sector may change to financial activities (sector K, in the CAE.rev 3 classification). Non-observed firms, due to the last two motives, may be observed in future years, if they pass quality control and/or return to a non-financial economic activity. Interesting also, is the fact that firm dynamics are dominated by entries (with only two years that see higher exits than entries, 2012 and 2014). From a dynamics point of view, we see that these percentages are relatively constant in terms of total firms reporting in the period. While stable in percentage, absolute values allow the observation of entry and exit spikes. These dynamics allows us to conclude for a relative stability in the behaviour of these firms.

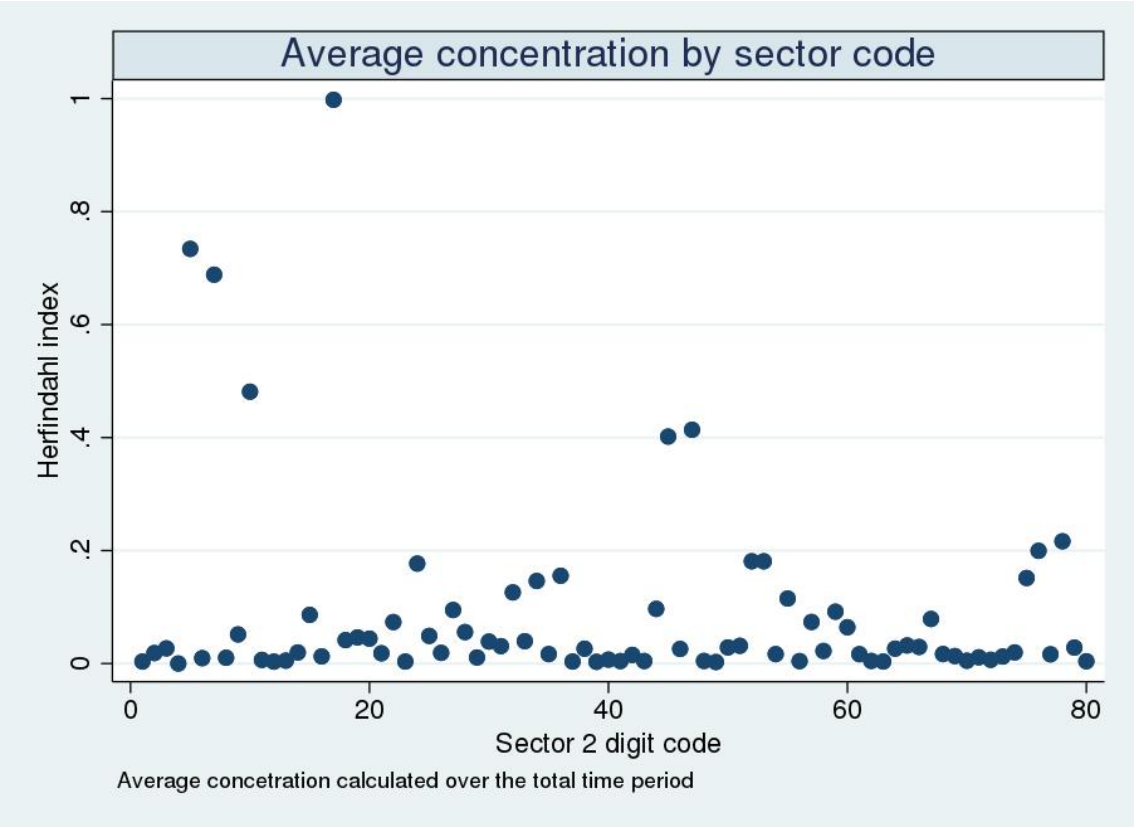
3.3. Sectorial behaviour

The database identifies the sector of operation of the company by using a detailed code of economic activity (5-digit CAE). Assuming that the activity code selected by the company reflects its actual economic activity, it is possible to calculate important statistics to understand the firm's position in the market. This subsection details the sector information in the database and shows the average concentration by sector.

Sector-wise (at two digit classification CAE rev 3), the database presents data on 98 sectors of economic activity for each year. With the exclusion of sector *06 – Oil & gas extraction* all sectors have at least 11 observations, with a minimum of one observation per year. The majority of the firms under analysis obtain their revenues from activities linked to their primary activity identifier representing almost 99% of the sample. The remaining 1.49% of firms, obtain their revenue from other activities linked to their non-primary activity code.[14] Different sectors have different dynamics and characteristics, and one such characteristic is market/sector concentration. The information on the sector of activity allows us to calculate a measure of market concentration that will give an idea about the existing level of competitiveness on a firm's market. Using as a measure of sector concentration the Herfindhal Index applied to the firms' market share, it is possible to see very different levels of concentration by sector, which are illustrated in the following graph.

[14] Appendix 2 shows the frequency table of output percentage derived from their first activity code.

Figure 1 - Average concentration by sector - whole time period



3.4. Firm size

Firm size is advanced as a relevant variable in the study of firm's performance and their export status. Firm size may be denoted by asset value, total sales, number of employees, or a combination of these metrics, reaching a final classification.[15] Given the other variables of interest, we considered firm size to be equal to the total assets of the firm. As the majority of the variables in the database, this variable has a large skew, with a minimum of zero and a maximum higher than a billion euros. Other than the size of the firm, its relative size to its sector of activity is analysed. This is done by calculating the market share of the largest firm in the sector (the leader) and comparing the market share of each firm to the leader's market share, in a yearly basis.[16]

[15] See for example EU recommendation 2003/36 on Small and Medium-sized Enterprises (SME) characterization.

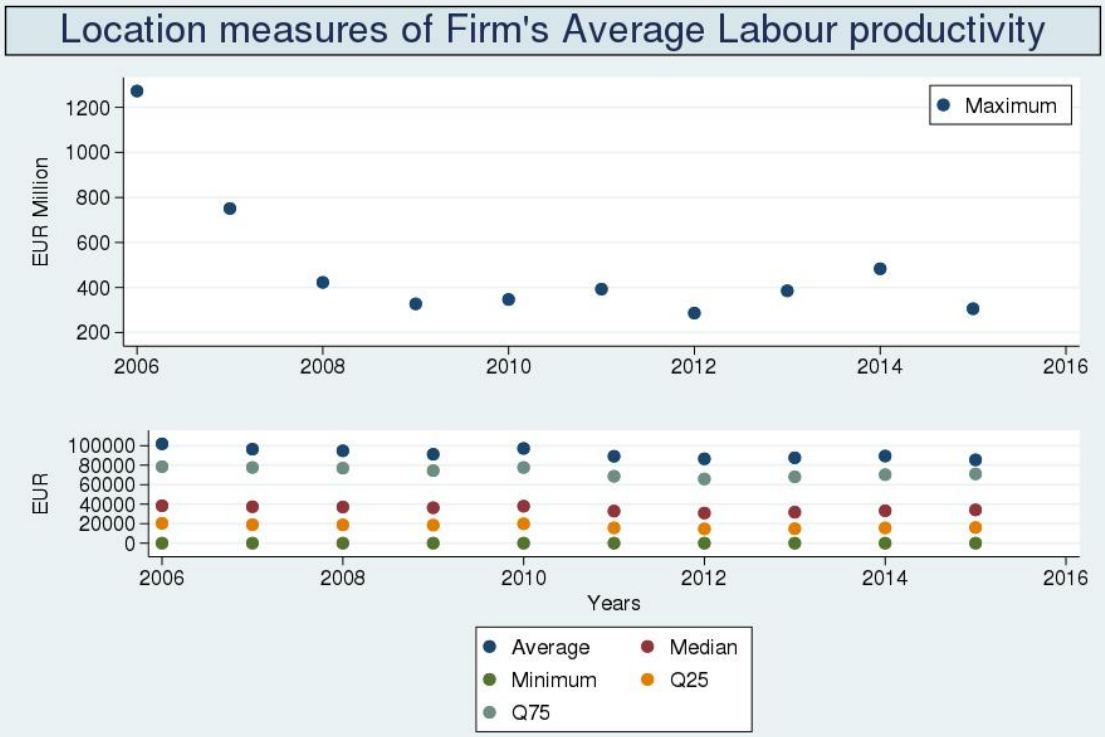
[16] Appendix 3 shows the absolute number of firms in each category of SME categorization.

3.5. Firm productivity

Firm productivity is an important aspect in the analysis of exporters. From one theory, self-selection, it is advocated that only the more productive firms start selling abroad. On the other hand, learning-by-exporting, defends that firms start improving their productivity after starting to export – *ex-post* phenomenon. This section defines the productivity proxy used in the study.

Two productivity measures were considered for the firm's average productivity, namely, output per worker and output per labour cost. From detailed statistics analysed, we concluded that output per labour cost was an inappropriate proxy for the firm's productivity, given that the variable presented extreme fluctuations on a yearly basis. In addition to the volatility of the variable, due to the fact that our analysis encompasses entities from various sectors, which can have different wage premiums, and wage growth rates, it was considered as a more reliable variable the use of output per worker. Given this decision, the analysis will be based on the average productivity by employee, given the more structural nature of employment versus labour costs in the sample. The following graphic shows the time evolution of the main location measures for the variable under use.

Figure 2 - Evolution of main location measures for average productivity by employee



The evolution of the firm's main location measures for productivity shows a stable temporal trend, while also presenting an extremely left concentrated distribution, as can be seen from the difference in scale when regarding quartiles and average productivity.[17] This result follows the pattern of most variables in the sample, and is in fact explained by the large number of very small firms (micro sized firms) in Portugal, which in addition to SME[18] account for almost all of the universe of firms in the country. Productivity follows this concentration as well, as the vast majority of firms (77% of total observations), present average productivity levels per employee below the one million euro mark, per year of average productivity.

[17] A boxplot of the variable, per year, was analysed, but not included due to the squashing of the main part of the graphic, the box, in value near zero.

[18] SME – Small and Medium Enterprises.

Table 4 - Frequencies of firm's productivity by categories of value

Productivity category	Freq.	Perc.	Cum.
<10k	407,099	10.97%	10.97%
10k-100k	1,973,336	53.19%	64.17%
100k-1M	477,482	12.87%	77.04%
1M-10M	17,245	0.46%	77.50%
10M-100M	788	0.02%	77.53%
100M-1kM	84	0.00%	77.53%
>1kM	1	0.00%	77.53%
Missing	833,648	22.47%	100.00%
Total	**3,709,683**	**100.00%**	**-**

3.6. Firm investment

The literature identifies capital per worker and investment as the principal variables influencing the decision process of a firm when assessing whether to become an exporter. Also, investment has been seen as a conscious effort performed by the firm in order to increase productivity and start exporting in the future. This section identifies and selects the proxy used in the study.

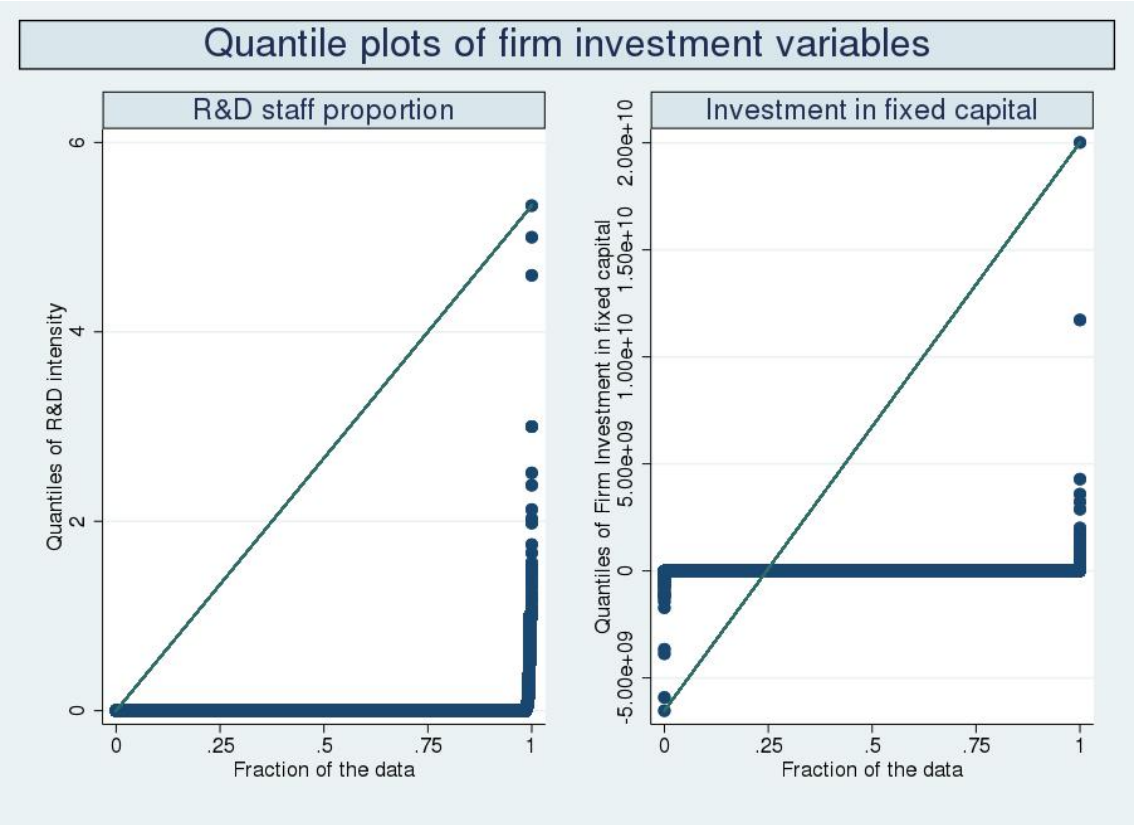
Firm's investment has been considered based in two different metrics. First, the investment in fixed capital, being this calculated as the net variation on non-current assets after adding asset depreciation cost. This metric allows to see the increase or decrease in a firm long term productive assets, combining tangible assets, such as, buildings, machines and land, intangible assets, such as trademarks, patents and brands, financial assets, encompassing the investment in other firms capital, and other non-current assets. By adding the cost derived from depreciation and amortization of assets, the metric gives, on a yearly basis, the net investment or divesture, in long term assets. Secondly, firms may also invest and develop Research and Development (R&D) activities, linked to the search for new or improved products, functionalities, processes or others. This investments may not have a monetary value associated in the balance sheet of companies, but may lead to productivity growth or positive impact on a firm's specific characteristics (such as, better management and/or processes). Given that R&D expenses were not of mandatory divulgation, this variable isn't directly observable. On the other hand, a large part of expenses linked to R&D are directly linked to the employees involved in these innovative tasks. Given this assumption, we define investment in R&D as proportion of employees dedicated to R&D on the total number of employees of the firm.

These variables show a high concentration near the zero value and while the investment in fixed capital shows a large dispersion, R&D staff proportion does not. Having a differentiated distribution, both these variables can provide different insights on firm's behaviour. Investment in fixed capital shows an expected distribution, centred at low positive values, with the existence of high investments and divestures during the period under analysis. For R&D staff proportion, the abnormal values (proportion above 1, or 100% of employed staff) are explained by the professional link of certain categories of employees to the firm (for example, interns or employees under scientific scholarships) who do not count for total employment statistics on the firm. Nonetheless, only 38 observation show these values.

Table 5 - Summary description for fixed capital investment and R&D staff proportion

Statistic	Fixed Capital Investment	R&D Investment
mean	85,500	0.1%
sd	17,720,922	0.0%
min	-6,531,892,736	0.0%
p25	0	0.0%
p50	325	0.0%
p75	9,361	0.0%
max	20,015,187,968	530.0%

Figure 3 - Quantile plots of firm investment variables

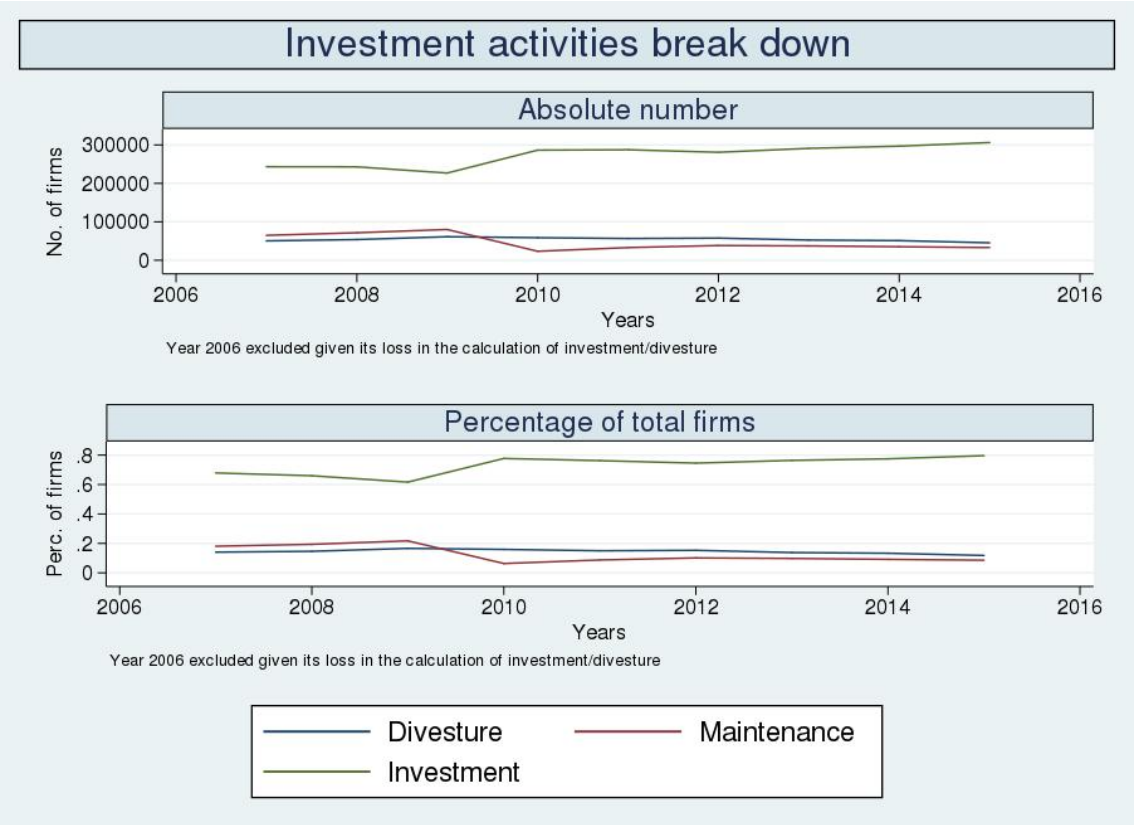


In order to further analyse the distribution of firm's investment in fixed capital, the following table shows, broken down per year, the number of firms, with existing value for the variable, reporting negative, neutral and positive investment. It is possible to see that, in absolute terms, the displayed values are relatively stable, evincing a positive trend in the number of firms investing. Nonetheless, comparing this to the plot of the percentage of firms in each activity as a proportion of the total number of reporting firms, different conclusions are reached. Two main trends are observable, namely, the stagnation of divesture firms and the increase in firms investing in 2009, which has stabilized in the following year until the end of the sample.

Table 6 - Breakdown of number of firm's investing, divesting or maintaining their asset structure by year

Period	Divesture	Maintenance	Investment
2007	50,399	64,745	243,355
2008	53,826	71,468	243,073
2009	61,215	79,957	226,947
2010	58,720	23,380	286,449
2011	56,607	32,936	287,471
2012	57,618	38,255	280,969
2013	52,517	37,244	291,138
2014	51,054	35,167	296,872
2015	45,375	32,968	306,335
Total	**487,331**	**416,120**	**2,462,609**

Figure 4 - Evolution of investment activities in absolute number and in percentage of total reporting firms



Overall, investment in fixed capital was preferred given the fact that it has a higher number of observations, and that the distribution is not truncated at 0.

Chapter 4 – Firm heterogeneity – the case of Portuguese exporters

This characterization will be based on the data base described in the previous section. Portuguese firms' demographics and characteristics are wildly disperse showing the existence of substantial heterogeneity. The use of detailed micro data allows for the identification of more detailed patterns than under a framework of representative agent or median agent.

A firm may be involved in international trade by either exporting, importing, receiving or sending direct foreign investment. The data allows us to study firm's activities in relation to the import and export of goods and services, in monetary cumulative values. From the point of view of the study, our focus will be on exports with a particular interest on the study of those characteristics that make a firm an exporter. This leads to the creation of two categories according to a threshold. We define as Exporter type A a firm that exports any positive amount while an Exporter type B, is a firm that exports more than 5% of its output. Exporters are also classified along other dimensions. First, firms may be classified as *entrants* or *exiters* in a given period if they start or stop exporting, while they are classified as *incumbent exporter* (*incumbent non-exporter*) if they maintain their exporter (non-exporter) status. Secondly, firms may be characterized with respect to the full period as *occasional* if they start and stop exporting, at least one time during the sample period, and as *permanent*, if they export for the full length of the period, or as *never exporter*, if the firm never exports during the period of analysis.

The consideration of two exporter types (A and B) according to exports thresholds allows us to analyse if there is a difference in behaviour between the different types of exporters. From the analysis of Figure 7 it is possible to identify a monotonous positive trend on the percentage of exporters with the exclusion of the decrease in 2009.[19] This trend is evident in absolute levels and percentage, as well as the similitude in the behaviour of both exporter types. This graphic leads to the conclusion that the extensive margin, e.g., the number of exporters, has increased in Portugal during the period.[20]

[19] The year is marked by many authors as the start of the Great Recession after the bankruptcy of Lehman Brothers.

[20] Appendix 3 presents the absolute and relative values of each type of exporters.

Figure 5 - Evolution of number and percentage of exporters

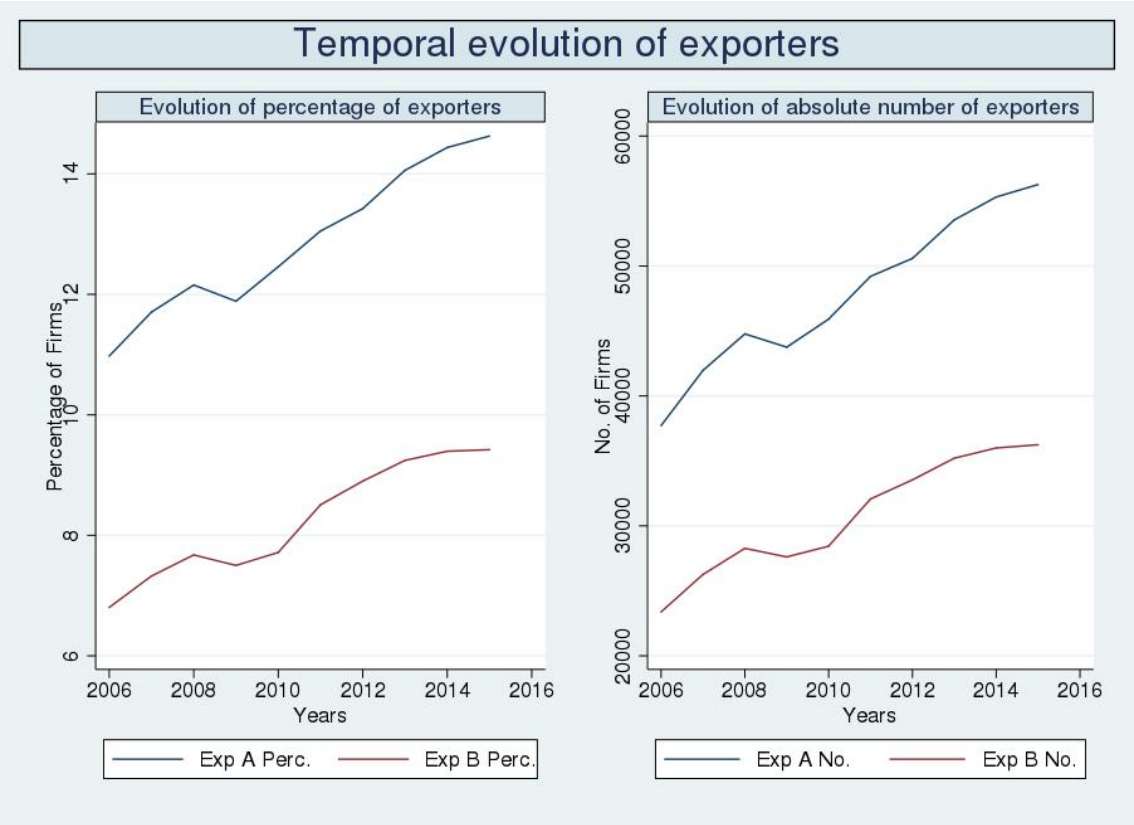
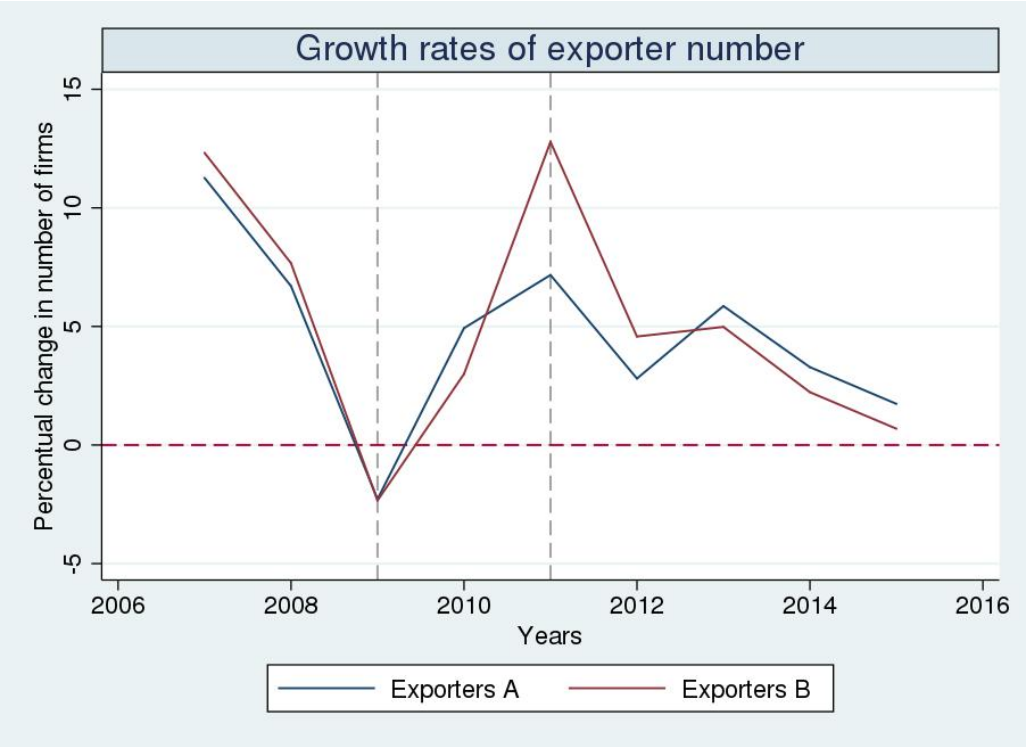


Figure 6 - Growth rate of exporter number, by type of exporter



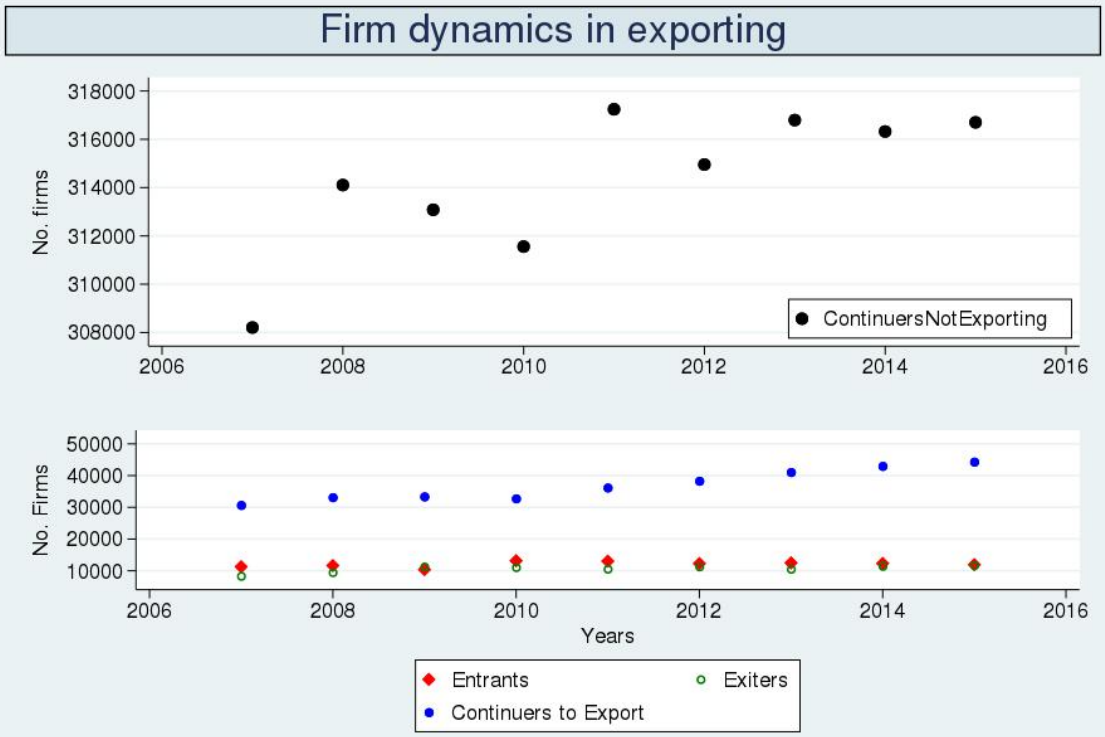
In terms of growth rates, three periods can be identified. The first, until 2009, where the growth rate of exporters decreased, reaching even a negative value, denoting the start of the financial crisis. The second, from 2009 to 2011, marks a period of increased growth rates, which can be attributed to the need for firms to start exporting due to the contraction in the internal market corroborating the study of López (2009). Finally, the third period marks a certain convergence towards zero. While both types of exporters show similar dynamics, for the period of 2010 to 2012, Exporters B show a more volatile growth rate, which can be explained in part due to the fluctuation around the cut-off value, representing noise in the sample.[21] Given the overall similitude in the behaviour of the variables, the following analysis will be based on Exporters type A, in order not to exclude from the category of exporters, occasional and small intensity exporters, and also, to avoid the noise created from the fluctuation of firm's exports around the cut-off value.

The overall dynamics of the number of exporters can be analysed by the decomposition of *entrants*, *exiters*, firms that are already exporters (*continuing exporting*) and firms that continue not to export (*continuing not exporting*). A certain stability can be seen in terms of *entrants* and *exiters*, with a slight dominance for *entrants*, which justifies the positive trend of *continuing exporting*, evidencing the presence of persistence in the exporters. The sum of these two effects, dictates the increase of exporters as shown above. In terms of non-exporters, while showing long term increase, they have seen severe fluctuations, showing that the majority of firm dynamics, in terms of entry and exit from economic activity is felt by non-exporters, which may be due to their lower resilience to economic fluctuations, linked to the dependence of the internal market.[22]

[21] Appendix 4 presents the yearly growth rates of each type of exporter.

[22] Appendix 5 shows the total number of *entrant*, *exiters*, *continuing exporters* and *non-exporters* per year

Figure 7 - Evolution of firm dynamics in exporting



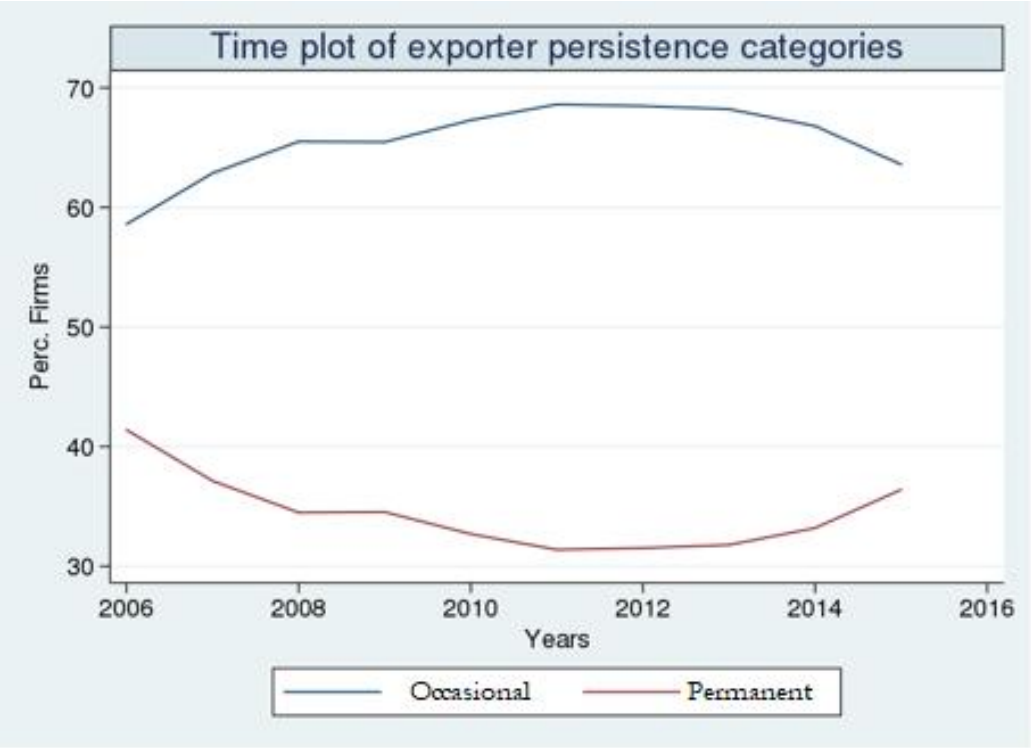
This dynamic, when considering the absolute number of firms involved in exporting activities and those that do not indulge in this activities, allow for the conclusion of the drop of the weight of non-exporters, from 89.0% to 85.4% on the total number of firms. This movement has been counterbalanced by the increase in Exporters over the decade.[23]

Other than the entry and exit of firms, an important characterization is that of the persistence of the export activity, namely, if the firm never exports, if the firm is an occasional exporter that exports, exits, and starts to export again at least once, or if the firm is a permanent exporter that exports for the full temporal range. During the period under analysis, total exporting firms have seen a slight increase in number as stated before, while their composition has seen a different behaviour. Occasional exporters have risen in the total number of exporters, from near 60% to almost 70% from 2006 until 2011, year where this trend inverts, and an increase in the percentage of permanent exporters increase throughout the remaining period. This increase may be due to the fact that as periods pass, a firm has to

[23] Appendix 6 shows the percentage of exporters and non-exporters in total firms by year.

be an exporter for less years to be considered as a permanent exporter. The dynamic is shown in the following graphic while in the appendix we report the actual values.

Figure 8 - Evolution of occasional and permanent exporters



Another dimension besides the number of exporters, or extensive margin, is the average quantity of exports each firm outputs, known as intensive margin, and as before, this shows a very concentrated distribution. The distribution shows two points of concentration, near 0% and 100%, showing that many firms, when exporting, direct almost their full production to the external market, representing the second largest category of exporters. Export intensity categories were created considering the average export intensity for each firm for the available period, being the top three classes [0%;10%],]90%, 100%] and]10%;20%], with shares of 79.59%, 12.85% and 2.41%, respectively.

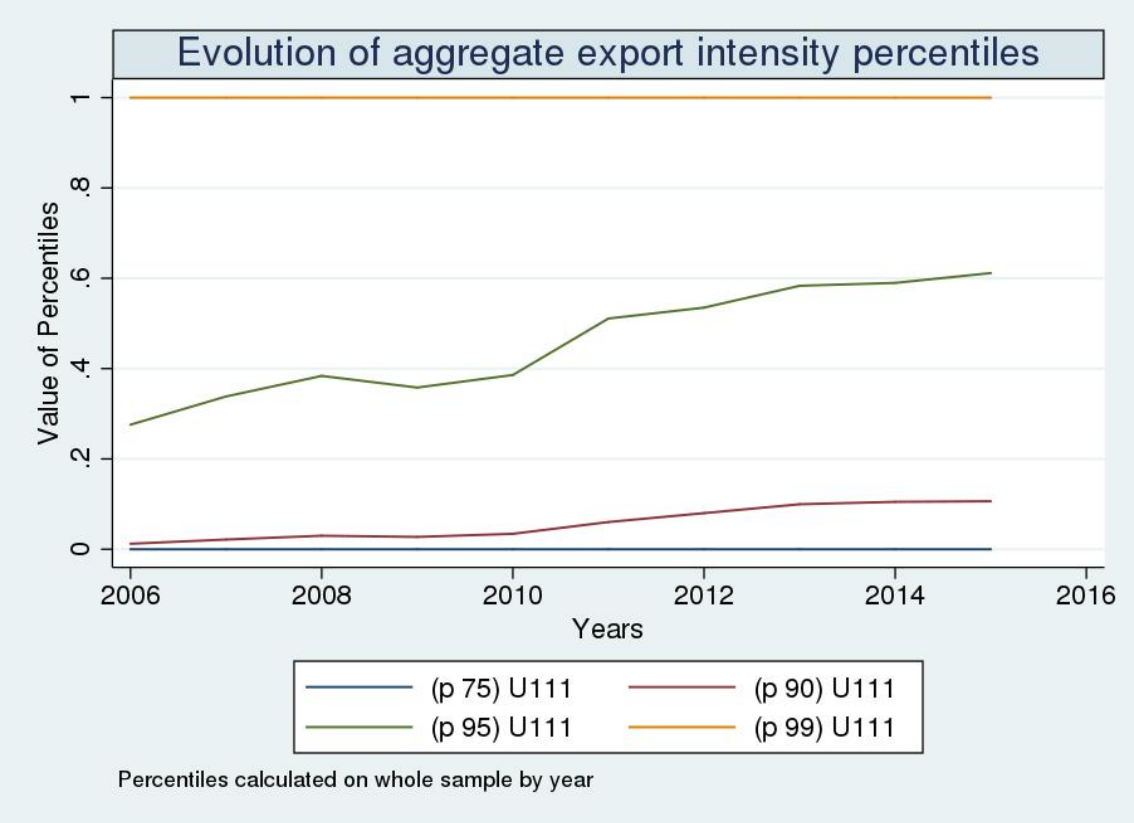
Table 7 - Frequency of export intensity by decile

Export intensity categories	Freq	Perc.	Cum.
[0%;10%]	48, 958	79.59%	79.59%
]10%;20%]	14,760	2.41%	82.00%
]20%;30%]	8,319	1.36%	83.35%
]30%;40%]	5,709	0.93%	84.28%
]40%;50%]	5,355	0.87%	85.16%
]50%;60%]	3,166	0.52%	85.67%

Export intensity categories	Freq	Perc.	Cum.
]60%;70%]	3,257	0.53%	86.20%
]70%;80%]	2,877	0.47%	86.67%
]80%;90%]	2,949	0.48%	87.15%
]90%;100%]	78,757	12.85%	100.00%
Total	**613,107**	**100.00%**	**-**

In dynamic terms, it can be seen a detachment of the overall distribution above the 75th percentile, namely, with the increase of the 90th and 95th percentile. The increase in the value of the 90th and 95th percentile shows that more exporters have been increasing their intensive margin, redirecting more of their production for external markets as time goes by.

Figure 9 - Percentile of export intensity temporal evolution



Chapter 5 - Methodology

5.1. Introduction

The objective of this thesis is twofold: first, to estimate and assess the existence of a productivity premium of exporters versus non-exporters thus testing the validity of the self-selection theory in Portuguese firms and, second, to study the determinants on the decision of a firm to become an exporter.

5.2. Part I – Assessing the existence of self-selection

5.2.1. Exporter productivity premium

Regarding the assessment of the self-selection theory, we follow the works of Wagner (2007) and Bernard and Jensen (1997), to check for the existence of an exporter premium for the Portuguese economy during the 2006-2015 period. The objective of this regression is to check for correlation between being an exporter and the productivity of the firm. This premium is calculated as the difference on labour productivity between exporters and non-exporters, after controlling for specific characteristics. For this estimation, we use the following equation:

$$Ln\ LP_{it} = \delta + \beta * Exporter_{it} + \gamma * Controls_{it}^{(1)} + \epsilon_{it} \qquad (5.1.)$$

where LP represents the labour productivity of firm *i* in year *t*, measured by the total sales of the firm divided by number of employees. Exporter is a dummy which takes the value *1* if the firm exports in that year. *Controls*$^{(1)}$ vector was used in order to isolate firm observable effects and include: 1) Sector, using a two digit dummy for the economic classification; 2) Region, using a district level dummy; 3) Category of firm size, using the Portuguese and European classification for SME firms. [24] Year dummies were also included in order to control for unobserved time effects, such as inflation or terms of trade.

For robustness, another equation was run, where firm level fixed effects were employed, as well as, time fixed effects. With the specification in equation 5.1. the firm fixed effects captures all observable and non-observable time-invariant characteristics of the firm that account for productivity differences, including the time invariant controls employed. Additionally, a firm's non-observable characteristics encompass a broad range of aspects

[24] This set of controls is designated "C*ontrols* $^{(1)}$"

such as different quality of the work force and human capital of workers, management quality, production processes, among others, which can be controlled for using fixed effects. Afterwards, the firm specific fixed effect is regressed on the exporter dummy, allowing one to conclude whether this intrinsic effect on productivity is larger for exporters when compared to non-exporters. This procedure allows for the analysis if the previous estimated effect is robust to an exhaustive set of controls, which tries to isolate the exporter effect on productivity from a large set of conditions. More formally, we first estimate a regression of type:

$$Ln\ LP_{it} = \hat{\alpha}_i + \hat{\alpha}_t + \hat{\gamma} * Controls_{it}^{(2)} + \epsilon_{it} \quad (5.2.)$$

followed by:

$$\alpha_i = \delta + \beta * Exporter_{it} + \epsilon_{it} \quad (5.3.)$$

The previous controls introduced in equation (5.1) being firm specific and time invariant are absorbed by the firm fixed effect. Thus, in (5.2) we use the following controls.[25] 1) Firm size, measured by the logarithm of the absolute value of total assets; 2) Firm profitability, calculated as the net return on shareholder's equity (return-on-equity, "ROE"); 3) Leverage, measured by the division of total debt by total shareholder's equity; 4) Firm investment, calculated as the division of net investment in fixed capital by the absolute value of firms asset;[26] 5) Relative size to sector, calculated as the division of the firm's market share by the market share of the leader of the sector, for each year;[27] 6) Market concentration, measured by the Herfindahl index for the sector, on a yearly basis; and 7) Market growth rate, measured as the growth rate of internal turnover/sales.[28]

After analysing the relation of being an exporter to the level of the firm's productivity, we modify the previous specification changing the dependent variable from the level of labour productivity, to its growth rate. This regression allows to assess if exporters

[25] This set of controls is designated "controls$^{(2)}$"

[26] Net investment is calculated as the variation of non-current assets added of the period depreciations.

[27] The leader of the sector is assumed to be the firm with the largest market share for the year. This value is assessed yearly.

[28] The controls were chosen in order to capture the effect of the firm's characteristics, capturing a large set of dimensions identified in the literature as affecting export propensity. Additionally, firm's leverage was added, extending from the dimensions identified in the literature. Appendix 7 sows the sources of the dimensions used as controls.

have higher labour productivity growth rates than non-exporters, for contemporaneous growth rates:

$$(Ln\ LP_{it} - Ln\ LP_{it-1}) = \delta + \beta * Exporter_{it} + \gamma * Controls_{it}^{(1)} + \epsilon_i \qquad (5.4.)$$

The controls introduced are the same as in equation 5.1. This equation tries to estimate the impact of being an exporter in the growth rate of labour productivity for the same year. This specification diverges slightly from Bernard and Jensen (1997) and Wagner (2007), as we consider more than two periods.

5.2.2. *Ex-ante* exporter productivity premium

In order to analyse if future exporters are already more productive than future non-exporters, the *ex-ante* productivity premium of exporters is calculated. The objective of this regression is to analyse if future exporters are today more productive than non-exporters, giving insights on the application of the self-selection theory on Portuguese firms. This is measured comparing the labour productivity of firms, subject to controls and to exporter status in the future. Thus, we run the regression

$$Ln\ LP_{it-z} = \delta + \beta * Exporter_{it} + \gamma * Controls_{it-z}^{(1)} + \epsilon_{it} \qquad (5.5)$$

The parameter z indicates the number of lags of analysis, between the export moment and the "basis" year. Different lags will be used allowing for the selection of the best fitting model. In order to assess the existence of this premium in terms of growth rate premium, we use a variation on equation 5.4., accounting for the lagged period as in equation 5.5.

$$(Ln\ LP_{it-z} - Ln\ LP_{it-z-1}) = \delta + \beta * Exporter_{i_t} + \gamma * Controls_{it-z}^{(1)} + \epsilon_i \quad (5.6)$$

The growth rates were calculated as the change in the logarithm of the labour productivity in the lag under test. The different values for the lag under use (by the change of z), allow for the analysis of a higher distance in the observation of the controls, and the event of exporting. This step diverges from the literature as we try to assess the best fitting period, instead of assuming a certain lag as the best fitting period.

5.3. Part II – Assessing the existence of learning-by-exporting

Another relevant aspect under analysis is the existence of possible learning effects, which may corroborate the theory of learning-by-exporting. In order to asses this phenomenon, the *ex-post* exporter productivity premium is estimated.

5.3.1. *Ex-post* exporter productivity premium

As before, we segment the analysis in two regressions, one in levels, and another one in growth rates in order to assess the existence of an *ex-post* exporter premium. The specification follows the one used for the inference of *ex-ante* productivity premium, but the time difference in periods (lags), are now changed to forward/future time periods.

$$Ln\ LP_{it+z} = \delta + \beta * Exporter_{it} + \gamma * Controls_{it+z}^{(1)} + \epsilon_{it} \qquad (5.7.)$$

The following specification for growth rate premium differs from that of The International Study Group on Exports and Productivity (2008) as the growth rates used are yearly growth rates and not multi-year growth rates.

$$(Ln\ LP_{it+z+1} - Ln\ LP_{it+z}) = \delta + \beta * Exporter_{i_t} + \gamma * Controls_{it+z}^{(1)} + \epsilon_i \quad (5.8)$$

5.3. Part II – Determinants of exporting

In a second part of this dissertation, the study focus turns to the determinants on the probability of a firm becoming an exporter.

This change in behaviour is modelled in the data as a binary variable, and as such, changes the scope of the model used from linear regression models to binary choice panel data models. A binary model is considered to be the appropriate choice for the case in analysis but the direct inclusion of firm fixed effects raises a problem because of the large number of variables included.[29] In addition to this computational problem there is the incidental parameter problem – the inconsistency and bias from estimating each individual fixed effect coefficient from T_i observations is likely to contaminate the remaining coefficients of the equation. This may be a problem because for each firm we have from 1 to 10 observations. Given the advanced considerations, two models could be used, namely a Random Effects Probit or Fixed Effects Logit.[30] Since there is the possibility of correlation between the explanatory variables and the error term (motivated by omitted variables), the assumption needed for the Random Effects model is not regarded as adequate, and as such,

[29] In the case, $(K + N)$ variables, corresponding to the number of coefficients of the explanatory variables and the firm specific coefficients. K represents the number of variables in the regression and N the number of observations.

[30] The estimation of a Fixed Effects Probit is not possible as the fixed effects have to be introduced explicitly requiring estimation of a model with a very large number of coefficients.

the Logit Fixed Effects model was chosen as the most appropriate model. This model is estimated with conditional maximum likelihood, under the following specification:[31]

$$EXP_{it}^{*} = \alpha_i + \alpha_{t-z} + \beta * Ln\ Prod_{it-z} + \gamma * Controls_{it-z}^{(2)} + \epsilon_{it} \quad (5.9)$$

$$EXP_{it} = 1\ if\ EXp_{it}^{*} > 0;$$

where EXP_{it}^{*} is the unobserved propensity to export, whose behaviour is modelled by observed explanatory variables (namely, the controls described on the previous subsection as *Controls*$^{(2)}$ (as used in equation 5.2), by a specific fixed firm effect and a fixed year effect. The observed dependent variable, whether a firm exports or not (Exp_{it}) is only observed when the unobserved variable is positive. The parameters α_i model the firm level effects, and cancel out of the estimation. The lag z is chosen based on the best fitting model in the estimation of the *ex-ante* exporter productivity premium. Finally, $Ln\ LP_{it-z}$ represents the measure of the firm's labour productivity, as used in the previous section.[32]

[31] Please see Greene (2012), Chapter 17, and Wooldridge (2002), Chapter 15 for further details on binary choice panel models.

[32] Please refer to Appendix 8 for the summary statistics of the variables used in the regressions.

Chapter 6 - Results

6.1. Introduction

This chapter is divided in two sections were specific results are presented. The first section assesses the existence of a productivity premium for exporters, both in the logarithm of the productivity and in the productivity growth rate. The *ex-ante* productivity premium of future exporters compared to future non-exporters, as well as the *ex-post* productivity premium of exporters compared to non-exporters is analysed in the same section. The second section, shows the results on the study of the determinants of exporting, with the use of a binary choice fixed effects model.

6.2. Assessing productivity premium for exporters

This section focuses on the identification of the productivity premium, either contemporaneous, *ex-ante* or *ex-post*. The possible endogeneity in the first set of regressions is acknowledged but we follow the methodology of other authors with adaptations. For comparison purposes, we also provide estimates for a regression without controls (baseline regression), allowing for the analysis of the impact of the variable of interest in the dependent variable without controlling for firm characteristics.

6.2.1. Exporter productivity level premium

The present subsection shows the results of the models implied by equation 5.1 and discusses their main findings.

Table 8 - Exporter productivity premium – equation 5.1.

VARIABLES	(1) Baseline	(2) Equation 5.1.
Exporter	0.760***	0.551***
	(0.00385)	(0.00353)
Firm Size Category = 2, Small		0.373***
		(0.00382)
Firm Size Category = 3, Medium		0.689***
		(0.0111)
Firm Size Category = 4, Big		0.988***
		(0.0328)
Constant	10.55***	10.46***
	(0.00220)	(0.0130)
Observations	2,729,203	2,723,489
Firm FE	NO	NO
Year FE	YES	YES
Sector	NO	YES
Region	NO	YES
Adj. Rsquared	0.0590	0.191

Note: Firm level cluster robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 10 presents the results from the regression of the logarithm of labour productivity on an exporter dummy (only variable in the baseline model), and several controls, as well as, a baseline regression without the referred dummies (the exception being year). From the estimates, it is possible to identify that the exporter variable is significant, leading to the conclusion that being an exporter has an impact on the labour productivity of the firm. The reported estimates of 114% ($e^{0.760} - 1$) and 73% ($e^{0.551} - 1$) premium are significant and indicate a high export premium. Even with the inclusion of controls for firm observable characteristics, the goodness-of-fit increases given the increase of the adjusted R-squared. According to Wagner (2007) these findings are corroborated by authors such as Colombia (1981-1991) or Isgut (2001) who found a positive correlation between exporter status and firm productivity, favouring the evidence that exporters are more productive than non-exporters. With the caveats noted regarding endogeneity, we can derive other conclusions, namely in terms of the significance of the time dummies and in terms of the impact of firm size. Although not reported, time dummies show negative estimates compared to the base year of 2006, which increase in magnitude in recent years, indicating that more recent observations show, ceteris paribus, lower productivity measures than in the

base year. Firm size also shows increasing magnitude with the increase on firm size, indicating that bigger firms show higher labour productivity than the base class of micro-sized firms.

In order to analyse the robustness of the productivity premium, we estimated equation 5.2. where a two-step methodology was applied. Step one consisted in regressing labour productivity on a vector of controls and firm and time fixed effects. Next we obtain estimates of the firm specific fixed effect. Then, on step two, the firm fixed effects estimates were run on a linear regression on an exporter dummy. This methodology allows for robustness check on the productivity premium found, as it estimates the productivity premium after including a large number of controls for both observable and unobservable firm characteristics.

Table 9 - Exporter premium fixed effects decomposition – equation 5.3.

	(1)
VARIABLES	Equation 5.3.
Exporter	0.243***
	(0.00354)
Constant	-0.0444***
	(0.00153)
Observations	2,046,799
Firm FE	NO
Year FE	NO
Controls	NO
Adj. Rsquared	0.0120

Note: Firm level cluster robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Note 2: The set of controls employed in the first step of the regression was the set controls(2)

The results in Table 11 show evidence of an exporter premium of 27.50% ($e^{0.760} - 1$), which while lower in magnitude to the previous estimates, show qualitatively the same result, that exporters present higher labour productivity. In terms of adjustment quality, the adjusted R-squared is 1.2%, which represents a low adjustment to the firm fixed effects estimate. Nonetheless, it is relevant to note that the estimated fixed effect encompasses the effect of all observable and non-observable time-invariant firm characteristics showing a clear difference between exporters and non-exporters.

These two sets of regressions allow for the conclusion that the findings regarding the existence of an exporter premium for Portuguese firms is robust to different methodologies. Even when controlling for several effects, being them observable or non-observable, the effect remains. Also, the estimates are clearly high, being located in a high end of the results found in the literature. This phenomenon may be caused by the inclusion in the sample of

small sized firms (in opposition to most other studies that only work with representative samples or only with a subset of the exporter population).

6.2.2. Exporter productivity growth rate premium

Following on the evidence of the existence of an exporter productivity premium, it is also relevant to study the existence of an advantage in the productivity growth rates of exporters. The literature advances that exporters are not only more productive, but they also show higher rates of productivity growth.

Table 10 - Exporter productivity premium in growth rates - equation 5.4.

VARIABLES	(1) Baseline	(2) Equation 5.4.
Exporter	0.0408***	0.0565***
	(0.000943)	(0.00114)
Firm Size Category = 2, Small		-0.0201***
		(0.000995)
Firm Size Category = 3, Medium		-0.0217***
		(0.00202)
Firm Size Category = 4, Big		-0.0185***
		(0.00434)
Constant	0.0207***	0.0869***
	(0.00128)	(0.00336)
Observations	2,197,679	2,192,654
Firm FE	NO	NO
Year FE	YES	YES
Sector	NO	YES
Region	NO	YES
Adj. Rsquared	0.00394	0.00549

Note: Firm level cluster robust standard error in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 12 shows the results of equation 5.4. Again, the exporter dummies are significant and show estimates of gains between 4 ($e^{0.0408} - 1$) and 6 ($e^{0.0565} - 1$) positive percentage points in the growth rate of labour productivity. Firm category size shows an expected size, with larger firms having lower productivity growth rates than smaller firms, which links with the previous findings of Cabral and Mata (2003). As larger firms show higher productivity than smaller firms, it is also expected that the productivity growth rate for these larger firms should be lower than the growth experienced by smaller firms. In terms of goodness-of-fit, the adjusted R-squared lowers to values below 1%, indicating that while the exporter dummies are significant, this accounts for a low explanatory power of the variability of productivity growth rate.

6.2.3. *Ex-ante* exporter productivity level premium

Other than considering the contemporaneous correlation between exporter status and labour productivity,[33] in order to assess the application of the self-selection theory to Portuguese firms, the *ex-ante* export premium is analysed, to see if future exporters are already more productive. For this objective, we estimate the *ex-ante* productivity premium for the productivity levels and its growth rate. To assess the impact of the productivity premium, different lags were used, and the adjusted R-squared was considered to choose the best fitting model.

Table 11 - *Ex-ante* exporter productivity premium - equation 5.5.

VARIABLES	(1) z=1	(2) z=2	(3) z=3	(4) z=4	(5) z=5
Exporter (t)	0.517***	0.493***	0.471***	0.448***	0.429***
	(0.00366)	(0.00383)	(0.00405)	(0.00428)	(0.00453)
Firm Size Category = Small (t-z)	0.355***	0.340***	0.324***	0.311***	0.299***
	(0.00392)	(0.00405)	(0.00420)	(0.00438)	(0.00460)
Firm Size Category = Medium (t-z)	0.682***	0.674***	0.666***	0.657***	0.646***
	(0.0114)	(0.0117)	(0.0121)	(0.0125)	(0.0131)
Firm Size Category = Big (t-z)	0.978***	0.968***	0.958***	0.950***	0.943***
	(0.0333)	(0.0337)	(0.0343)	(0.0349)	(0.0356)
Constant	10.45***	10.44***	10.44***	10.42***	10.40***
	(0.0135)	(0.0141)	(0.0148)	(0.0154)	(0.0161)
Observations	2,325,158	1,950,151	1,608,680	1,301,937	1,022,470
Firm FE	NO	NO	NO	NO	NO
Year FE	YES	YES	YES	YES	YES
Sector	YES	YES	YES	YES	YES
Region	YES	YES	YES	YES	YES
Adj. Rsquared	0.198	0.203	0.208	0.211	0.216

Note: Firm level cluster robust standard error in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 13 shows the estimates for five different lags, where all variables except the exporter dummy are lagged to the number of periods specified by z. The exporter dummy remains significant in all regressions, as well as the controls for firm size. The estimates for the exporter dummy decrease in a monotonous way, indicating a stabilization around a 56% ($e^{0.0408} - 1$) *ex-ante* exporter productivity premium level. Firm size dummies show similar behaviour as in previous estimations, increasing with firm size, and maintaining stable levels as lags are increased, due to being in the same time period that the dependent variable. The

[33] The contemporaneous regressions were run as tested in the literature with the consideration that endogeneity may arise. For a discussion on the cause and/or effect of exporting on a firm's performance please see Bernard and Jensen (1997).

loss in the number of observations and number of firms is justified due to use of increasing lags, eliminating the first time periods from the sample. Overall, the adjusted R-squared increases with the lags used, being maximum (for the considered time periods) at z=5 at 21.6%. The decrease in the *ex-ante* exporter premium show a divergence in the productivity levels of future exporters compared to future non-exporters, which linked to the existence of an exporter productivity premium in the furthest lag may indicate the existence of possible learning effects, where productive firms accumulate further productivity (linked to their higher productivity growth rates). These findings show that Portuguese future exporters are more productive than future non-exporters. Our specification diverges from the selected authors in the use of several time lags, and our results show higher magnitude than in the literature. Nonetheless, quantitatively, our results corroborate other authors' findings, and provide evidence for the self-selection theory.

6.2.4. *Ex-ante* exporter productivity growth rate premium

As in the contemporaneous setting, in order to assess the existence of the premium under study, a regression is performed on the growth rate of labour productivity in order to estimate if future exporters see higher productivity growth rates than non-exporters. These estimates indicate if future exporters already have higher productivity growth rates, or if this increase is due to being an exporter.

Table 12 - *Ex-ante* exporter productivity premium in growth rate - equation 5.6.

	(1)	(2)	(3)	(4)	(5)
VARIABLES	z=1	z=2	z=3	z=4	z=5
Exporter	0.0576***	0.0542***	0.0481***	0.0418***	0.0369***
	(0.00116)	(0.00125)	(0.00134)	(0.00147)	(0.00168)
Firm Size Category = 2,	-0.0292***	-0.0274***	-0.0262***	-0.0249***	-0.0281***
	(0.00104)	(0.00112)	(0.00121)	(0.00133)	(0.00153)
Firm Size Category = 3,	-0.0304***	-0.0274***	-0.0242***	-0.0205***	-0.0268***
	(0.00213)	(0.00230)	(0.00250)	(0.00282)	(0.00323)
Firm Size Category = 4,	-0.0257***	-0.0255***	-0.0212***	-0.0222***	-0.0300***
	(0.00442)	(0.00463)	(0.00510)	(0.00547)	(0.00655)
Constant	0.0872***	0.0865***	0.0943***	0.0737***	0.0515***
	(0.00351)	(0.00375)	(0.00403)	(0.00451)	(0.00509)
Observations	1,852,553	1,534,498	1,246,488	982,039	739,990
Firm FE	NO	NO	NO	NO	NO
Year FE	YES	YES	YES	YES	YES
Sector	YES	YES	YES	YES	YES
Region	YES	YES	YES	YES	YES
Adj. Rsquared	0.00581	0.00565	0.00593	0.00532	0.00374

Note: Firm level cluster robust standard error in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The table above shows the estimates for the exporter dummy and several controls in the firm's labour productivity, with all variable except for the exporter dummy are lagged z periods. As in Table 13 the estimates for the exporter premium decrease with the lags while remaining significant. The controls follow the same behaviour. On the other side, firm size shows a different behaviour than when contemporaneous estimation was run, with the estimates remaining similar independently of firm size. Sample size is further penalized with the use of growth rates, leading to the loss of one additional period in each regression due to the calculation of growth rates with two observations at a time. Overall, while adjusted R-squared levels are similar, below 1%, this metric assumes its local maximum at z=3. As before, the tendency in terms of exporter premium show a divergence between future exporters and future non-exporters.

The main conclusions of these subchapters focusing on *ex-ante* exporter premium are for the existence of a productivity premium both in absolute productivity level and in productivity growth rates, which decreases as lags increase in the regression. These findings corroborate the self-selection theory for Portuguese firms during the study period, given than future exporters show favourable characteristics in terms of productivity compared to non-exporters. Nonetheless, these findings do not guarantee causality, meaning that it is not possible to confirm that it is the higher productivity level that leads firms to start to export. This question is analysed in section 6.3. resorting to another methodology.

6.2.5. *Ex-post* exporter productivity level premium

The *ex-post* exporter productivity premium allows for the investigation if exporting causes increases in productivity, meaning, if firm's productivity increases as a consequence of being an exporter.

Table 13 - *Ex-post* exporter productivity premium - equation 5.7.

VARIABLES	(1) z=1	(2) z=2	(3) z=3	(4) z=4	(5) z=5
Exporter (t-z)	0.493***	0.460***	0.438***	0.428***	0.420***
	(0.00374)	(0.00402)	(0.00435)	(0.00473)	(0.00517)
Firm Size Category = Small (t)	0.360***	0.374***	0.392***	0.411***	0.427***
	(0.00396)	(0.00417)	(0.00443)	(0.00472)	(0.00504)
Firm Size Category = Medium (t)	0.682***	0.703***	0.728***	0.753***	0.773***
	(0.0114)	(0.0118)	(0.0124)	(0.0129)	(0.0135)
Firm Size Category = Big (t)	0.976***	0.985***	1.000***	1.017***	1.030***
	(0.0333)	(0.0340)	(0.0347)	(0.0355)	(0.0365)
Constant	10.49***	10.53***	10.52***	10.54***	10.51***
	(0.0134)	(0.0141)	(0.0148)	(0.0157)	(0.0166)

Observations	2,305,498	1,896,349	1,539,057	1,231,199	954,954
Firm FE	NO	NO	NO	NO	NO
Year FE	YES	YES	YES	YES	YES
Sector	YES	YES	YES	YES	YES
Region	YES	YES	YES	YES	YES
Adj. Rsquared	0.194	0.196	0.198	0.201	0.205

Note: Firm level cluster robust standard error in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 15 shows significant coefficients for the exporter dummy, indicating that the exporter status in the past, has a significant a positive effect on the firm's productivity today. As in previous results, the estimates decrease as time lags increase, showing that while estimates are positive, this relation decreases in magnitude, indicating possible decreasing marginal returns (corroborating the findings of Martins and Yang (2009)). Also, firm size dummies are significant and show the same qualitative behaviour as before, increasing in magnitude when the exporter dummy increases in lag. In terms of R-squares, it peaks at z=5, reaching close to 21% of explanatory power.

6.2.6. *Ex-post* exporter productivity growth rate premium

Following the same methodology as in the *ex-ante* exporter premium estimation, the dependent variable is changed to the labour productivity growth rate, in order to assess if becoming an exporter, increases the growth rate of firms in future periods.

Table 14 - *Ex-post* exporter productivity premium in growth rate - equation 5.8.

VARIABLES	(1) z=1	(2) z=2	(3) z=3	(4) z=4	(5) z=5
Exporter (t-z)	-0.0296***	-0.0107***	-0.00494***	0.00211	-5.01e-05
	(0.00117)	(0.00120)	(0.00132)	(0.00148)	(0.00168)
Firm Size Category = Small (t)	-0.0031***	0.0244***	0.0267***	0.0270***	0.0269***
	(0.00100)	(0.00101)	(0.00108)	(0.00118)	(0.00130)
Firm Size Category = Medium (t)	0.0136***	0.0438***	0.0441***	0.0448***	0.0448***
	(0.00202)	(0.00205)	(0.00211)	(0.00227)	(0.00244)
Firm Size Category = Big (t)	0.0251***	0.0543***	0.0596***	0.0606***	0.0556***
	(0.00417)	(0.00402)	(0.00422)	(0.00440)	(0.00471)
Constant	0.0960***	0.0476***	-0.0102***	0.0361***	-0.0391***
	(0.00337)	(0.00338)	(0.00362)	(0.00392)	(0.00440)
Observations	2,192,654	1,834,214	1,492,645	1,195,785	931,243
Firm FE	NO	NO	NO	NO	NO
Year FE	YES	YES	YES	YES	YES
Sector	YES	YES	YES	YES	YES
Region	YES	YES	YES	YES	YES
Adj. Rsquared	0.00490	0.00590	0.00640	0.00704	0.00785

Note: Firm level cluster robust standard error in parentheses. *** p<0.01, ** p<0.05, * p<0.1

While before the exporter variable presented positive estimates, this is altered in Table 16 with the indication of negative estimates for the exporter dummy. This results being

significant, indicate that being an exporter is linked to having future lower growth rates. The estimates for higher lags lose significance without losing explanatory power. These findings are in line with the literature, with several authors encountering negative or insignificant effects when estimating *ex-post* exporter premiums. Firm size estimates also present positive coefficients, which may be linked to the contemporaneous nature of the variables. In terms of goodness-of-fit, the model with z=5 presents the highest R-squared, while this is below 1% as in previous regressions.

6.3. Exporting determinants

The previous section has presented evidence on the existence of exporter premium, in the several time periods analysed (contemporaneous, *ex-ante* and *ex-post*). After considering this evidence, it is clear that exporters present higher labour productivity than non-exporters, finding correlation between exporting and productivity. In order to access causality, meaning, estimating if being more productive, increases the chances of a firm being an exporter, we estimate a binary choice model with firm labour productivity as an explanatory variable. The use of such a model also allows for the study of the impact of several other variables in the event of exporting. This section focuses on the effect of a set of control variables and of productivity in the decision of a firm being an exporter, giving valuable insight to policymakers on what dimensions of the firm to influence in order to increase exporter numbers in the economy.

In this section a non-linear model is used in order to assess and estimate the impact of some variables (characteristics of the firm) on the probability that a firm is considered an exporter, with particular focus on the effect of the labour productivity.

Table 15 - Export determinants estimation – equation 5.9

VARIABLES	(1) z=1	(2) z=3	(3) z=5
Firm labour productivity (t-z)	0.246***	0.0463***	0.0325*
	(0.00811)	(0.0113)	(0.0181)
Firm absolute size (t-z)	0.453***	0.229***	0.0656***
	(0.0110)	(0.0155)	(0.0253)
Firm return on equity (t-z)	2.16e-05	-2.66e-05	2.83e-05
	(1.70e-05)	(1.89e-05)	(5.11e-05)
Debt on shareholder equity ratio (t-z)	5.64e-06	-1.84e-05	6.90e-06
	(6.42e-06)	(1.72e-05)	(2.50e-05)
Relative investment (t-z)	0.149***	0.0513*	0.0150
	(0.0234)	(0.0282)	(0.0455)
Firm relative output in sector (t-z)	1.915***	1.323***	0.512

VARIABLES	(1) z=1	(2) z=3	(3) z=5
	(0.338)	(0.447)	(0.748)
Sector concentration (t-z)	0.619***	-0.0473	0.304
	(0.176)	(0.234)	(0.290)
GR of sector internal sales (t-z)	-0.000297	0.00190	0.00336
	(0.000581)	(0.00156)	(0.00233)
Observations	401,322	232,470	103,766
Number of firms	63,599	44,777	27,895
Firm FE	YES	YES	YES
Year FE	YES	YES	YES
Pseudo Rsquared	0.0255	0.00385	0.000766

Note: Firm level cluster robust standard error in parentheses. *** p < 0.01 , ** p < 0.05, * p< 0.1
Note2: The set of controls employed in the regression was the set controls[2]

The regression results show interesting insights on the variables used, which allow for short term and long term considerations. Labour productivity shows significance and a positive estimate for the entire period of analysis, while lowering its significance at z=5. These findings show that labour productivity has a positive effect on the probability of a firm being an exporter, corroborating the self-selection theory. Effectively, from the results shown, it is possible to infer that higher productivity firms, show higher chances of being exporters, corroborating a certain causality relation in the variables.

This behaviour is followed by firm absolute size, as firms with higher asset values, show evidence of having higher probability of being exporters. This links with the findings of other variables, namely, firm's relative size to the sector also with positive significant estimates (until lag z=3). This effect may be due to the fact that as firm size increases (either in absolute or relative size), the relative size of the outstanding market decreases and, as such, a firm's growth potential is limited, unless it starts operating in external markets.

These two variables show significance at several periods, indicating a medium long effect on the export probability, meaning, that asset accumulation and productivity have a long term effect on the probability of being an exporter. Increasing a firm's assets and/or productivity, impacts its export propensity several years from the event. Also, this effect is also felt in the short term, as an increase in this variables, impacts the probability of being an exporter in the following period. This short term relation may indicate that firms increase their asset base or productivity in the eminence of starting to export.

Profitability, leverage and growth rate of internal sales do not show significance, indicating that there is no statistical evidence of their impact in the probability of exporter. This is specifically important in the case of the behaviour of internal sales, as this counters

previous findings in the literature that firms start exporting when their internal sales drop, in order to use their fixed productive capacity (as evidenced by Blum et al. (2013)).

Two other variables show a significant impact in exporter propensity, but only for the short and medium term. The positive estimate for lag χ=1 of relative investment, links to the conclusions of the previous paragraphs and the literature, as firms invest in order to allow them to start exporting (corroborating the hypothesis of López (2009)). In the same tendency, sector concentration also seems to increase competitive markets, tend to have higher probabilities of being exporters. The reason for this behaviour may be that firms avoid the increasing competition in the market, or that due to the competitive pressure, they acquired sufficient productivity levels that allow them to surpass entry barriers to exporting.

The smaller sample sizes compared to the previous sections are due to the fact that firms that do not change exporting status during the period do not contribute to the estimation. Nonetheless the model without doubt shows the existence of a positive effect of lagged productivity on the probability of a firm becoming an exporter. As such, these estimates seem to confirm the existence of a causal effect, between the increase of productivity and the increase in the probability of becoming an exporter, in future periods, corroborating the self-selection theory.

In order to test the robustness of the results of the impact of labour productivity, the model was run with the controls lagged only at one period, isolating labour productivity in its several lags from the influence of the controls.[34] The results obtained were similar. The main differences are the increase in significance of "Firm Return on Equity", "Debt on Shareholder Equity Ratio" and "Firm relative output in the sector" lagged one period show significance. Additionally, "Sector Concentration" also shows significance when labour productivity is lagged three or five periods. These changes in significance are mainly due to the fact that these variables are lagged only one period, lessening the possibility of part of their effect being absorbed by the productivity variable. Also, this new specification while increasing the goodness-of-fit, decreases the coefficient magnitude associated with labour productivity.

[34] This regression is presented in Appendix 9.

Chapter 7 – Conclusion

This project proposed to study the behaviour of exporter firms. First, we checked for the existence of a productivity premiums for exporters. Secondly, in order to assess the application of the self-selection and learning-by-exporting theories to the reality of Portuguese firms, we estimated the *ex-ante* and *ex-post* exporter productivity premium. Finally, the causality of productivity on the probability of becoming an exporter was explored with the use of a logit fixed effects model.

The main findings of this project were the existence of a contemporaneous exporter premium, in level and growth rate of productivity, suggesting that exporters have higher productivity than non-exporters. These findings are robust to different methodologies, as was verified with the estimation of the firm's fixed effects and the use of an exporter dummy in a second-step regression. Afterwards, in order to assess the application of the self-selection theory, we estimated the *ex-ante* exporter premium, finding that future exporters have *ex-ante* higher productivity than future non-exporters. This seems to indicate that future exporters are already more productive, and show higher productivity growth rates. Finally, in order to assess the effect of being an exporter in the productivity of the firm, the *ex-post* exporter productivity premium was estimated. We found robust evidence for the existence of *ex-post* premiums in levels, indicating that exporters show higher future productivity levels than non-exporters. This effect is not observed when estimating the impact of being an exporter in the future productivity growth rate of the firm, as only the first three period show significance. The estimates for the three first lagged periods, show negative coefficients, indicating that exporters evidence in the following years lower productivity growth rates than non-exporters, which may be motivated by the adjustments and investments needed after entering external markets.

Lastly, to further analyse the causality relation between productivity and exporting, in an attempt to study if more productive firms are more likely to be exporters, a panel binary model with fixed effects was employed. This analysis shows that an increase in productivity has a positive and significant impact on the probability of being an exporter. This effect is followed by firm absolute size in terms of long term effects. In terms of medium and short term impacts, relative investment and firm's relative output to the sector leader show a positive impact until a three period lag, and sector concentration shows positive estimates only at a one period lag.

In terms of economic policy, the implications are relevant given that future exporters have higher productivity levels and growth rates, being capable of generating more added-value than their non-exporter counterparties. The fact that productivity has an impact on exporting, means that if policymakers want to increase the extensive margin, e.g. the number of firms exporting, incentives to productivity should have a positive impact in the agent goal. Also, higher concentration sectors and bigger relative firms are more likely to be exporters. Nonetheless, given the negative effects, ceteris paribus, of more concentrated sectors and bigger relative firms (indicating of concentration), policymakers should focus on increasing the number of exporters via the increase in individual firms' productivity. Lastly, in terms of relative investment, this may generate positive effects, both in terms of firm's efficiency and of innovation, as such, a directed action to increase the investment of a firm relative to its absolute size may be a sound action by the policymaker in order to increase the number of exporters in the economy.

Future avenues of research could explore the robustness of our finding with respect to the applicability of the self-selection theory to the Portuguese firms. One possibility is the inclusion of the lagged exporter dummy in the binary model, leading to a dynamic panel study. Another possibility is the use of matched sample techniques as proposed by several authors in order to reduce the possibility of bias due to sample selection issues. Yet another, is the use of alternative productivity measures, such as TFP. While interesting, these endeavours are well beyond the scope of this work.

Bibliography

Alvarez, R., & López, R. A. (2005). Exporting and performance: evidence from Chilean plants. *Canadian Journal of Economics/Revue canadienne d'économique, 38*(4), 1384-1400. doi: doi:10.1111/j.0008-4085.2005.00329.x

Alvarez, R., & López, R. A. (2008). Entry and Exit in International Markets: Evidence from Chilean Data*. *Review of International Economics, 16*(4), 692-708. doi: 10.1111/j.1467-9396.2008.00749.x

Ball, L. M. (2014). Long-Term Damage from the Great Recession in OECD Countries. *NBER Working Paper Series, 20185*. doi:10.3386/w20185

Banco de Portugal. (2017). *Central Balance Sheet Database - Data Manual.*

Bernard, A. B., & Jensen, J. B. (1997). Exceptional Exporter Performance: Cause, Effect, or Both? *National Bureau of Economic Research Working Paper Series, No. 6272*. doi: 10.3386/w6272

Bernard, A. B., & Jensen, J. B. (2001). Why Some Firms Export. *National Bureau of Economic Research Working Paper Series, No. 8349*. doi: 10.3386/w8349

Bernard, A. B., & Jensen, J. B. (2004). Exporting and Productivity in the USA. *Oxford Review of Economic Policy, 20*(3), 343-357. doi: 10.1093/oxrep/grh020

Blum, B. S., Claro, S., & Horstmann, I. J. (2013). Occasional and perennial exporters. *Journal of International Economics, 90*(1), 65-74. doi: https://doi.org/10.1016/j.jinteco.2012.11.002

Burda, M., & Wyplosz, C. (2013). *Macroeconomics - A European Text*: Oxford University Press.

Cabral, L. M. B., & Mata, J. (2003). On the Evolution of the Firm Size Distribution: Facts and Theory. *American Economic Review, 93*(4), 1075-1090. doi: doi: 10.1257/000282803769206205

Greene, W. H. (2012). *Econometric Analysis* (Seventh edition ed.): Pearson Education.

Hallward-Driemeier, M., Iarossi, G., & Sokoloff, K. L. (2002). Exports and Manufacturing Productivity in East Asia: A Comparative Analysis with Firm-Level Data. *National Bureau of Economic Research Working Paper Series, No. 8894*. doi: 10.3386/w8894

Impullitti , G., Irarrazabal, A. A., & Opromolla, L. D. (2012). A theory of entry into and exit from export markets. *Journal of International Economics*. doi: http://dx.doi.org/10.1016/j.jinteco.2012.11.005

JaeBin, A., & Alexander, M. (2013). Capacity Constrained Exporters: Micro Evidence and Macro Implications: Florida International University, Department of Economics.

Krugman, P. (1980). Scale Economies, Product Differentiation, and the Pattern of Trade. *American Economic Review, 70*(5), 950-959.

Lee, S.-H., Beamish, P. W., Lee, H.-U., & Park, J.-H. (2009). Strategic choice during economic crisis: Domestic market position, organizational capabilities and export flexibility. *Journal of World Business, 44*(1), 1-15. doi: https://doi.org/10.1016/j.jwb.2008.03.015

López, R. A. (2009). Do Firms Increase Productivity in Order to Become Exporters?*. *Oxford Bulletin of Economics and Statistics, 71*(5), 621-642. doi: 10.1111/j.1468-0084.2009.00558.x

Martins, P., & Yang, Y. (2009). The impact of exporting on firm productivity: a meta-analysis of the learning-by-exporting hypothesis. *Review of World Economics (Weltwirtschaftliches Archiv), 145*(3), 431-445.

Mayer, T., & Ottaviano, G. I. P. (2008). The Happy Few: The Internationalisation of European Firms. *Intereconomics, 43*(3), 135-148. doi: 10.1007/s10272-008-0247-x

Melitz. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica, 71*(6), 1695-1725.

Melitz, & Redding. (2014). Heterogeneous Firms and Trade *Handbook of International Economics, 4th ed* (Vol. 4, pp. 1-54): Elsevier.

Muñoz-Sepúlveda, J. A. (2014). *Residual exports and domestic demand: an empirical analysis*. MPRA Paper. University Library of Munich, Germany. Retrieved from https://ideas.repec.org/p/pra/mprapa/58328.html

Oliveira, M. R. (2016). *Bases de Microdados com Informação Económico Financeira da IES (Informação Empresarial Simplificada).* (Master), Faculdade de Economia da Universidade do Porto.

Silva, A., Afonso, O., & Africano, A. P. (2012). Learning-by-Exporting: What We Know and What We Would Like to Know. *The International Trade Journal, 26*(3), 255-288. doi: 10.1080/08853908.2012.682022

Sousa, J. (2016). *The impact of services export on economic growth: the Portuguese case.* (Master in International Business), Faculdade de Economia da Universidade do Porto (FEP.UP).

The International Study Group on Exports and Productivity. (2008). Exports and productivity – comparable evidence for 14 countries. (pp. 78). NBB Working Papers No. 128, 13 February 2008.

Timoshenko, O. A. (2015). Learning versus sunk costs explanations of export persistence. *European Economic Review, 79*(Complete), 113-128. doi: 10.1016/j.euroecorev.2015.02.006

Vagos, C. (2015). *A internacionalização da Economia Portuguesa via promoção das exportações e criação de valor acrescentado.* (Mestrado em Economia), Faculdade de Economia da Universidade do Porto (FEP.UP), FEP - Relatório de Estágio. Retrieved from http://hdl.handle.net/10216/80039

Wagner, J. (2007). Exports and Productivity: A Survey of the Evidence from Firm-level Data. *World Economy, 30*(1), 60-82. doi: 10.1111/j.1467-9701.2007.00872.x

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. London, England: The MIT Press.

Appendices

Appendix 1. User defined variables

Table 16 - User defined variables

Variable identifier	Variable Name	Calculation formula	Notes
U100	Labour productivity 1	Turnover / No. of employees (year end)	Missing values coded to zero
U101	Labour productivity 2	Turnover / Staff costs	Missing values coded to zero
U102	Firm size	Total Assets	The absolute value of total assets was considered, due to non-normal accounting values
U104	Firm profitability 1 (ROE)	Net Income / Shareholder's equity	Shareholder's realized equity was used for consistency
U106	Leverage ratio 2 (Debt-to-equity)	Total liabilities / Shareholder's equity	Shareholder's realized equity was used for consistency
U107	Exports (external sales)	Sales of goods and services to other territories	-
U108	Internal sales	Sales of goods and services in Portugal	-
U111	Export intensity	Exports / (U107 + U108)	The denominator sums the total turnover of the firm
U113	R&D Intensity	No. of employees (in R&D) / No. of employees (year end)	Percentage of employees linked to R&D activities
U114	Firm investment (fixed capital)	D.(Non-current Assets) + Depreciation and Amortisation costs	Represents the variation in non current assets not linked with capital depreciation
U116	Insignificant firms	True if Turnover < 1000 & Assets < 1000 & Employes = 0 OR Constitution year > 2015	Characterizes non-active firms from and economic point of view, considering that firms that report turnover and assets below a thousand euros and have no staff are "ghost" firms. Also, firms reporting their constitution year after the year of 2015 (year of data) where considered to be prne to reporting errors and were disregarded
U201_2lvl	CAE - 2 digits	-	Represents the sector of activity in terms of Group (Statistics Portugal methodology)
U206	Firm level market share - 2 digits	(U107 + U108) / U203	Firm total output divided by sector output
U207	Firm position in sector - 2 digits	U206 / leader-market-share	Firm market share divided by the sector leader market share, per year
U209	Concentration index (Herfindahl Index)	sum (U206^2)	Sum of individual firms squared market share, by year, by sector
U301	Exporter dummy - A	U107 > 0 ; recode (missing = 0)	Creates dummy with 1 if firm has export higher than zero, and 0 otherwise
U401	Exporter dummy - B	U107 > 5%*(U107 + U108) ; recode (missing = 0)	Creates dummy with 1 if firm has export higher than 5% total output, and 0 otherwise

Appendix 2. Output percentage derived from first activity code

Table 17 - Output percentage derived from first activity code

Percentile	Freq.	Percent	Cum.
1	453	0.07%	0.07%
2	88	0.01%	0.09%
3	148	0.02%	0.11%
4	237	0.04%	0.15%
5	600	0.10%	0.25%
6	666	0.11%	0.36%
7	1,212	0.20%	0.56%
8	2,045	0.33%	0.89%
9	3,676	0.60%	1.49%
10	603,982	98.51%	100.00%
	613,107	**100.00%**	-

Appendix 3. Yearly number and percentage of each exporter type

Table 18 - Number and percentage of exporter types

Year	Freq_Exp A	Perc_ Exp A	Freq_Exp B	Perc_Exp B	No. Firms
2006	37,716	11.00%	23,378	6.80%	343,623
2007	41,964	11.70%	26,258	7.30%	358,499
2008	44,776	12.20%	28,271	7.70%	368,367
2009	43,756	11.90%	27,610	7.50%	368,119
2010	45,913	12.50%	28,437	7.70%	368,549
2011	49,205	13.10%	32,072	8.50%	377,014
2012	50,585	13.40%	33,540	8.90%	376,842
2013	53,550	14.10%	35,211	9.20%	380,899
2014	55,309	14.40%	35,996	9.40%	383,093
2015	56,269	14.60%	36,244	9.40%	384,678
Total	**479,043**	**12.89%**	**307,017**	**8.24%**	**3,709,683**

Appendix 4. Yearly growth rate of each exporter type

Table 19 - Yearly growth rate of Exporter types

Year	Growth Rate - Exp A	Growth Rate - Exp B
2006	-	-
2007	11.30%	12.30%
2008	6.70%	7.70%
2009	-2.30%	-2.30%
2010	4.90%	3.00%
2011	7.20%	12.80%
2012	2.80%	4.60%
2013	5.90%	5.00%
2014	3.30%	2.20%
2015	1.70%	0.70%
Average	**4.61%**	**5.11%**

Appendix 5. Number of entrants, exiters, active exporters and non-exporters

Table 20 - Decomposition of entry, exit and maintenance variable per year

Year	Entrants	Exiters	Cont. Exporting	Cont. NOT Exporting
2006	0	0	37,716	305,907
2007	11,268	8,327	30,696	308,208
2008	11,647	9,484	33,129	314,107
2009	10,383	11,280	33,373	313,083
2010	13,148	11,076	32,765	311,560
2011	13,024	10,565	36,181	317,244
2012	12,281	11,302	38,304	314,955
2013	12,486	10,553	41,064	316,796
2014	12,316	11,460	42,993	316,324
2015	11,931	11,707	44,338	316,702
Total	**108,484**	**95,754**	**370,559**	**3,134,886**

Appendix 6. Percentage of exporters and non-exporters in total number of firms

Table 21 - Percentage of exporter and non-exporters in total number of firms per year

Year	Exporters	Non-exporters
2006	11.00%	89.00%
2007	11.70%	88.30%
2008	12.20%	87.80%
2009	11.90%	88.10%
2010	12.50%	87.50%
2011	13.10%	86.90%
2012	13.40%	86.60%
2013	14.10%	85.90%
2014	14.40%	85.60%
2015	14.60%	85.40%
Average	**12.89%**	**87.11%**

Appendix 7. Controls(2) vector variables selection

Table 22 - Control variables selection

Controls	Author
Productivity	Melitz (2003)
Firm size	Blum et al. (2013)
Profitability	Lee et al. (2009)
Relative investment	Alvarez and López (2005)
Relative size in sector	Lee et al. (2009)
Market concentration	Lee et al. (2009)
Market Internal growth rate	Blum et al. (2013)

Appendix 8. Summary statistics of relevant regression variables

Appendix 8.a. Number of firms per sector – exporter and non-exporters

Table 23 - Number of firms per sector - exporter and non-exporters

Sector 2 dígits	Total	Non-exporter	Exporter
1	92,358	83,898	8,460
2	13,427	12,114	1,313
3	5,829	4,789	1,040
6	5	5	0
7	224	198	26
8	8,892	6,636	2,256
9	204	164	40
10	56,055	48,177	7,878
11	8,283	4,734	3,549
12	43	22	21
13	20,809	12,707	8,102
14	47,222	31,064	16,158
15	19,646	12,619	7,027
16	29,357	21,803	7,554
17	4,272	2,321	1,951
18	20,933	14,947	5,986
19	163	128	35
20	7,045	4,372	2,673
21	1,491	967	524
22	9,922	4,835	5,087
23	27,877	17,559	10,318
24	2,886	1,636	1,250
25	66,175	45,560	20,615
26	2,496	1,447	1,049
27	5,594	3,253	2,341
28	12,935	7,415	5,520
29	4,536	2,337	2,199
30	1,957	1,111	846
31	24,809	16,438	8,371
32	14,319	10,892	3,427
33	17,844	13,475	4,369
35	7,203	6,861	342
36	1,406	1,301	105
37	540	469	71
38	7,368	5,631	1,737
39	190	170	20
41	285,898	270,996	14,902
42	28,704	25,737	2,967
43	150,277	134,374	15,903
45	142,895	123,186	19,709
46	362,284	254,848	107,436
47	497,998	458,182	39,816
49	168,145	139,154	28,991
50	2,767	1,998	769
51	738	441	297

Sector 2 dígits	Total	Non-exporter	Exporter
52	18,832	13,148	5,684
53	3,099	2,655	444
55	47,878	46,437	1,441
56	283,816	281,163	2,653
58	15,956	12,027	3,929
59	12,050	9,346	2,704
60	2,971	2,570	401
61	4,670	3,923	747
62	43,600	32,224	11,376
63	5,189	4,099	1,090
68	256,506	252,693	3,813
69	105,713	98,651	7,062
70	87,779	75,202	12,577
71	79,322	68,160	11,162
72	3,624	2,941	683
73	29,793	23,251	6,542
74	34,826	27,465	7,361
75	8,533	8,280	253
77	17,025	15,213	1,812
78	5,241	3,935	1,306
79	13,374	10,889	2,485
80	3,163	2,747	416
81	22,668	21,853	815
82	61,349	51,877	9,472
85	49,213	47,507	1,706
86	162,216	160,196	2,020
87	9,133	9,088	45
88	8,311	8,251	60
90	11,531	10,126	1,405
91	817	770	47
92	1,288	1,268	20
93	32,257	30,383	1,874
94	3,401	3,263	138
95	9,019	7,988	1,031
96	71,499	70,080	1,419
Number of sectors: 80			

Appendix 8.b. Number of firms per district – exporter and non-exporters

Table 24 - Number of firms per district - exporter and non-exporters

District	Total	Non-exporter	Exporter
Aveiro	235,825	193,192	42,633
Beja	39,214	35,777	3,437
Braga	277,106	223,516	53,590
Braganga	33,508	30,012	3,496
Castelo Branco	53,442	47,792	5,650
Coimbra	128,987	115,562	13,425
Ivora	52,506	46,834	5,672
Faro	181,734	171,281	10,453
Guarda	40,916	35,371	5,545
Leiria	190,710	159,188	31,522
Lisboa	1,053,718	932,633	121,085
Portalegre	31,023	26,788	4,235
Porto	649,700	548,533	101,167
Santarém	143,762	127,020	16,742
Setúbal	227,213	207,760	19,453
Viana do Castelo	66,419	53,604	12,815
Vila Real	48,711	43,079	5,632
Viseu	99,510	86,808	12,702
Angra do Heroismo	11,663	11,328	335
Horta	6,325	6,222	103
Ponta Delgada	25,951	25,326	625
Funchal	111,453	102,763	8,690
Desconhecido	287	251	36

Appendix 8.c. Number of firms per district – exporter and non-exporters

Table 25 - Number of firms per category size - exporter and non-exporters

Size category	Total	Non-exporters	Exporters
Micro	3,196,029	2,894,083	301,946
Small	382,016	248,245	133,771
Medium	59,589	24,634	34,955
Big	6,512	3,311	3,201

Appendix 8.d. Summary statistics of Controls[2] variables – exporter and non-exporters

Table 26 – Summary statistics of control[2] variables – exporter and non-exporters

Variable	Category	Obs	Mean	Std.Dev.	Min	Max
Firm labour productivity	Total	2,729,203	10.55	1.203	-6.551	20.96
	Non-exporter	2,273,659	10.42	1.179	-6.551	20.00
	Exporter	455,544	11.17	1.120	-5.298	20.96
Firm absolute size	Total	3,688,536	11.61	1.989	-4.605	24.02
	Non-exporter	3,210,051	11.40	1.924	-4.605	23.02
	Exporter	478,485	12.99	1.856	-4.605	24.02
Firm return on equity	Total	3,658,726	6.14	16.67	-5,275,658	27,588,232
	Non-exporter	3,183,842	1.00	16.50	-5,275,658	27,588,232
	Exporter	474,884	40.64	17.78	-1,828,778	11,939,054
Debt on shareholder equity ratio	Total	3,592,928	308.7	104,938	-2,220,012	170,156,272
	Non-exporter	3,118,695	265.2	56,658	-2,220,012	59,856,052
	Exporter	474,233	594.4	249,639	-13,115	170,156,272
Relative investment	Total	2,490,685	-131	1.019e+06	-1.230e+09	231.5
	Non-exporter	2,100,843	-155	1.109e+06	-1.230e+09	231.5
	Exporter	389,842	-0.63	233.6	-109,783	18.83
Firm relative output in sector	Total	3,709,678	0.00316	0.0256	0.00	1.00
	Non-exporter	3,230,635	0.00183	0.0159	0.00	1.00
	Exporter	479,043	0.0122	0.0574	0.00	1.00
Sector concentration	Total	3,709,683	0.0150	0.0379	0.00	1.00
	Non-exporter	3,230,640	0.0147	0.0380	0.00	1.00
	Exporter	479,043	0.0171	0.0375	0.00147	1.00
Growth rate of internal sales	Total	3,080,409	0.0526	12.04	-1.00	14,294
	Non-exporter	2,660,928	0.0465	12.89	-1.00	14,294
	Exporter	419,481	0.0914	3.398	-1.00	895.60

Appendix 9. Binary model with controls lagged at one period

Table 27 - Export determinants estimation controls lagged at one period – equation 5.9.

VARIABLES	(1) z=1	(2) z=3	(3) z=5
Firm productivity (t-z)	0.246***	0.0339***	0.0247*
	(0.00811)	(0.00902)	(0.0138)
Firm absolute size (t-1)	0.453***	0.628***	0.622***
	(0.0110)	(0.0136)	(0.0222)
Firm Return on Equity (t-1)	2.16e-05	1.73e-05	0.00109**
	(1.70e-05)	(2.99e-05)	(0.000479)
Debt on Shareholder Equity Ratio (t-1)	5.64e-06	-1.10e-05	-0.000548***
	(6.42e-06)	(1.98e-05)	(0.000170)
Relative investment (t-1)	0.149***	0.00854	4.48e-05
	(0.0234)	(0.0141)	(0.00177)
Firm relative output in sector (t-1)	1.915***	2.441***	2.525***
	(0.338)	(0.406)	(0.645)
Sector concentration (t-1)	0.619***	1.046***	0.884**
	(0.176)	(0.275)	(0.374)
GR of sector internal sales (t-1)	-0.000297	-0.000743	-0.000318
	(0.000581)	(0.00143)	(0.000541)
Observations	401,322	298,266	145,466
Number of firms	63,599	51,908	32,749
Firm FE	YES	YES	YES
Year FE	YES	YES	YES
Pseudo Rsquared	0.0255	0.0190	0.00969

Note: Firm level cluster robust standard errors in parentheses. *** p < 0.01 , ** p < 0.05, * p< 0.1