Universitat Autònoma de Barcelona

The European Integration Process:
Trade, Mobility, and Policy

Alicia Gómez Tello

Universitat Autònoma de Barcelona
Departament d'Economia Aplicada

The European Integration Process: Trade, Mobility, and Policy

Alicia Gómez Tello

Supervised by Rosella Nicolini

Dissertation submitted to the
Department of Applied Economics
at the

UNIVERSITAT AUTÒNOMA DE BARCELONA

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in the subject of
Applied Economics

February 2016

To my parents, my sisters, and Javi.

Contents

Acknowledgments

The completion of this dissertation means the end of an unforgettable stage of my life. For this reason, I would like to express my gratitude to those people that have accompanied me during this adventure.

First, I am particularly grateful to my thesis director Rosella Nicolini for her dedication, continuous suggestions, encouragement, and support. Along these five years I learnt a lot from her and I will never forget her advice. I am truly indebted to her!

The unconditional support from Javi has also been crucial to accomplish this thesis. I really appreciate his understanding and patience. He always believed in my abilities and encouraged me in the moments of doubts.

Nothing would have been possible without the encouragement of my mother and my sisters Amparo and Ana. Maybe, they did not understand very well what I was doing, but they always supported me, above all in the most difficult stages. I would like also to thank the rest of my family (especially my aunt Nati, my uncles Boro and Frederic, and my cousins Salva and Àngela), Javi's family, and my friend Jorge.

During the time I spent at the Departament d'Economia Aplicada I had the opportunity to carry out my studies in a fair climate, with an open dialogue with other students and researchers. I would like to thank all the members of the Department, especially Professor Xavi Ramos for his valuable comments in all my seminars, my mates and office's partners Adriana, Ana, Andrés, David, Darío, John Fredy, Lidia, Luciana, Macarena, Mónica, Paola, Pedro, Roberto, Sayuri, Sonia, and Vania.

I am also very thankful to the members of the Institut de Recherches Économiques et Sociales (Belgium), for hosting me and the excellent research atmosphere they were able to create during my stay. It was a great experience, not only from a professional viewpoint but also from a personal viewpoint. I will always remember the stimulating discussions with Andrea Ciani and Sotiris Blanas.

I would like to acknowledge the full and step-in members of my thesis jury professors Miren Lafourcade, José Luis Roig Sabaté, Florian Mayneris, Ernest Miguélez Sánchez, and Kristian Estevez, for being able to take part of this evaluation process: I will definitely enjoy their expertise and suggestions.

Finally, the financial support from ECO2014-52506-R and 2014-SGR-327 is gratefully acknowledged.

Alicia Gómez Tello

February 2016

List of Acronyms

AEA	Anuario de Estadística Agraria
ATA	Agencia de Trabajadores Autónomos
F&B	Food and Beverages
BvD	Bureau van Dijk
CAP	Common Agricultural Policy
CD	Cobb-Douglas
CEE	Central and Eastern Europe
CEECs	Central and Eastern European Countries
CLC	CORINE Land Cover
CES	Constant Elasticity of Substitution
CMEA	Council of Mutual Economic Assistance (also COMECON)
CMOs	Common Market Organizations
CNAE	Clasificación Nacional de Actividades Económicas
CORINE	Coordination of Information of the Environment
DID	Difference-in-Differences
DVs	Dummy Variables
EAs	European Agreements
EC	European Commission
ECSC	European Coal and Steal Community
EEA	European Environmental Agency
EEC	European Economic Community
EEC-6	Refers to the six founding members of the EEC (i.e., Belgium, France, Germany, Italy, Luxembourg and the Netherlands).

EFTA	European Free Trade Association
EU	European Union
EU-15	Refers to the EU members from 1995 to 2004 (i.e., Austria, Belgium, Denmark, Finland, France, Greece, Germany, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden, and the United Kingdom).
EU-10	Refers to the 10 countries that joined the EU in 2004 (i.e., Cyprus, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia, and Slovenia).
EU-27	Refers to the EU members from 2007 to 2013 (i.e., without Croatia).
FAO	Food and Agricultural Organization of the United Nations
FAOSTAT	Food and Agricultural Organization of the United Nations Statistics Division
FDI	Foreign Direct Investment
FE	Fixed-Effects
FSU	Former Soviet Union
FOCs	First Order Conditions
GAEZ	Global Agro–ecological Zones
GDP	Gross Domestic Product
HH index	Herfindahl-Hirschman index
IIASA	International Institute for Applied Systems Analysis
IMF	International Monetary Fund
INE	Instituto Nacional de Estadística
IPF	Identificador de Persona Física
Ivie	Instituto Valenciano de Investigaciones Económicas
MCVL	Muestra Continua de Vidas Laborales
MAGRAMA	Ministerio de Agricultural, Alimentación y Medio Ambiente
MINECO	Ministerio de Economía y Competitividad
ML	Maximum-Likelihood
MNEs	Multinational enterprises
NACE	Nomenclature générale des Activités Économiques dans les Communautés Européennes

NB	Negative Binomial
NUTS	Nomenclature of Territorial Units for Statistics
OECD	Organisation of Economic and Co-operation Development
OLS	Ordinary Least Squares
PCAs	Partnership and Cooperation Agreements
PPC	Potential Production Capacity
PTAs	Preferential Trade Agreements
PWT	Penn World Table
ROW	Rest of World
SABI	Sistema de Análisis de Balances Ibéricos
SEs	Standard Errors
SIA	Sistema Integrado de Información de Agua
SITC	Standard International Trade Classification
TPC	Total Production Capacity
UNCTAD	United Nations Conference on Trade and Development
USSR	Union of Soviet Socialist Republics

List of Figures

List of Tables

Chapter 1

Introduction

To elucidate what European integration means today, it is critical to review the catastrophic socioeconomic situation in Europe directly following World War II (1939–1945),[1] when avoiding further international conflict ranked among Europe's top concern. However, in the postwar context, Europe's eastern and western parts were ideologically divided, and each proposed starkly different solutions to the situation affecting all nations and Europe as a whole. Whereas Western European countries sought stronger integration among all European nations, Eastern European countries, heavily influenced by the political and economic leadership of the Soviet Union, instead suggested the solution of a socialist planned economy.

Amid this social, political, and economic landscape, a sharp debate among countries of the Western Bloc sought to clarify how the proposed integration of European nations would be implemented. The group of countries supporting supranationalism—Austria, France, Luxembourg, the Netherlands, and Germany—advocated the creation of a federalist structure, in which part of the powers traditionally reserved for nations would be delegated to new supranational institutions. The other group of countries—Denmark, Iceland, Ireland, Norway, Sweden, Switzerland, and the United Kingdom—instead endorsed intergovernmentalism, which would minimize the reach of new supranational institutions and the influence of common policies. This somewhat antifederalist faction strongly supported the independence of European nations and sought limit cooperation among them to selected scopes.[2]

The first major step toward European federalism was taken in 1952, with the implementa-

[1]World War II left 15.6 million soldiers and 19.5 million civilians dead, 50 million people without home, cities and town in ruins, and water, heating, electricity, and transport infrastructure destroyed (Staab, 2013).

[2]Two models for intergovernmentalism were the Organisation for European Economic Cooperation in 1948, which evolved today's Organisation for Economic Co-operation and Development, and the Council of Europe created in 1949.

tion of the Schuman Plan inspired by Jean Monnet. French Foreign Minister Robert Schuman proposed the creation of a supranational authority that would control the production of coal and steel in France and Germany, chiefly in order to avoid another Franco–German military conflict, Six countries—Belgium, France, Germany, Italy, Luxembourg, and the Netherlands—joined the pact by signing the Treaty of Paris in 1951, and the European Coal and Steel Community (ECSC) was created the following year. Although the ECSC was a successful supranational institutions, it involved only the economic activity of two specific sectors. An additional measure of economic integration emerged in 1957, when the Treaty of Rome created the European Economic Community (EEC) and the European Atomic Energy (Euratom).[3] For the supranational cooperation of the EEC, three communitarian institutions were created: the European Parliament Assembly, the European Court of Justice, and the European Commission. Unlike the ECSC, the EEC was a true customs union involving all economic sectors.[4] The EEC additionally pursued the free movement of services, labor, and capital, the informal coordination of exchanges rates, and a less clearly defined commitment to common agricultural policy. Of course, most of those ends took several years to take complete effects. In the meantime, European nations less in favor of a federalist structure responded by creating the European Free Trade Association (EFTA) in 1960.[5]

During the next two decades, however, no progress occurred in the European integration process. Both negative political and economic conditions worldwide and the priority of protecting domestic industries generally thwarted the introduction of policies that would have fostered integration. Nevertheless, in 1973, three EFTA members— Austria, Ireland, and the United Kingdom—joined the EEC, citing purely economic reasons. Most notably, the economic performance of EEC countries was far superior to that of EFTA members. In fact, during 1950–1970, the gross domestic product of the six nations more than doubled that of the seven EFTA nations, and EEC incomes were growing twice as fast (Baldwin and Wyplosz, 2012).[6]

During 1985–1995, European Commission President Jacques Delors introduced the most

[3]The Treaty of Rome was ratified by the six countries that signed the Treaty of Paris in 1951—Belgium, France, Germany, Italy, Luxembourg, and the Netherlands. Also known as the Six, these nations are considered to be the founders of the European Union. Throughout this paper, we refer to these six countries as the EEC-6 or EU-6.

[4]The first step in an economic integration process—that is, the free trade agreement—was bypassed. Any free trade agreement guarantees only the free movement of goods among member states, whereas a customs union upholds the same trade policies toward third-party countries as well.

[5]The EFTA constituted a free trade agreement among Austria, Denmark, Norway, Portugal, Sweden, Switzerland, and the United Kingdom and involved all economic sectors except the agricultural production.

[6]The British government was the first to react to this large difference in the economic performance between the EEC and EFTA countries and in 1961 applied for EEC membership, but its application was not well perceived by France.

important measures toward true economic integration since the 1950s. Upon the approval of the Single European Act (1986), promises made by the Treaty of Rome were finally kept. For one, the act eased the creation of a single market for goods, services, capital and labor. At the same time, important changes in the functioning of European institutions facilitated progress toward a more integrated political and economic community. The Council of the EU applied qualify majority voting to take new decisions, for instance, thereby making it more difficult for a single country to veto proposed legislation.

Another important agreement during this period and a pillar of the European integration process was the Treaty of the European Union (i.e., the EU Treaty), signed in signed in Maastricht in 1992.[7] The EU Treaty created a monetary union and introduced measures to achieve a truly political union—for instance, European citizenship—as well as policies addressing common foreign and internal affairs. After the ratification of the EU Treaty, the EEC was officially renamed as European Union. Both before and after this period, the enlargement process continued; Greece joined in 1981, Portugal and Spain in 1986, and Austria, Finland, and Sweden in 1995. In 1995, the European Union was composed of 15 members—hence, the EU-15—and the number of member states remained unchanged until 2004.

At the end of the 1980s, European socialist economies initiated a *glasnost* process, which induced the transformation of their planned economies into free market economies. By and large, this process coincided with significant political turmoil. In 1990, Estonia, Latvia, and Lithuania declared their independence from the Soviet Union, which finally disintegrated in 1991. During the early 1990s, ten Central and Eastern European Countries (CEECs) applied for EU membership. Yet, due to major socioeconomic differences between the CEECs and EU-15 countries, it was necessary to implement a trial period in order to prepare member candidates for integration. The process mostly consisted of establishing and achieving milestones in shifting their economies from planned to market ones, as well as consolidating their democratic institutions. In May 2004, ten countries—Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia, and Slovenia—joined the European Union (i.e., the EU-10). To fulfill similar membership requirement, however, Bulgaria and Romania had to wait three more years.

In the 2000s, the European Union was a far larger institution than that created a half-century earlier and therefore was required to face new challenges in meeting the top goal of improving the standard of living of all citizens, as established by the Treaty of Rome. To that end, all

[7] The EU Treaty took force in November 1993, after a difficult ratification process in Denmark.

EU member states agreed upon a newly defined agenda eventually formalized with the Lisbon Treaty in 2007,[8] which introduced a range of new policies aimed to make the European Union a more democratic, efficient organization. Some specific goals related to employment level, gender equality, environment protection, and the quality of life of European citizens—that is, targets that had often been neglected. In fact, these principles provide the basis for the research and innovation program currently implemented by the European Union under the framework of the Horizon 2020 program.

At present, the European Union counts 28 members (after the entry of Croatia in 2013). Social, cultural, and economic differences remain important among member states. If cultural differences can be appreciated as distinguishing signs of identity, then economic imbalances could limit the positive effects associated with the integration process.

From this perspective, the objective of this dissertation is to examine the benefits and drawbacks of a number of selected features of the European integration process summarized above. Our research is relevant in the context of policy, since understanding the limits of EU policies from an economic viewpoint can afford insights into and suggestions for overcoming current obstacles, as well as for conceiving strategies for a more effective integration process. In the spirit of evaluating the integration process, this dissertation addresses three distinct aspects: trade integration among EU member states after the fifth EU enlargement (2004), the free mobility of capital and workers across EU member states, and the true effectiveness of the Common Agricultural Policy (CAP) as a pillar of the European communitarian policies.

Among our means toward those ends, in Chapter 2 we assess the extent to which the fifth EU enlargement induced the effective economic integration of new EU member states, by checking if they made EU-15 countries their privileged trade partners at the expense of other commercial partners in order to fulfill the economic predictions of a canonical customs union agreement. By extension, in Chapter 3 we address the true effectiveness of the integration process by referring to the degree of mobility among firms and workers within the European Union, with Spain as our point of reference. Lastly, in Chapter 4 we focus on analyzing the effectiveness of one of the European policies aiming at improving the integration in terms of production. Since its creation, the European Union has financed and implemented specific actions to improve the efficiency and productivity of economic activities across members. Among the most controversial and complex policies has been that concerning agriculture. In this dissertation, we therefore aim to bring

[8]The Lisbon Treaty entered into force in December 2009, following a difficult ratification process in Ireland.

new evidence to this discussion by developing a study that assess the effectiveness of the CAP in the case of the Spanish economy.

In what follows, we more precisely outline the content of the three core chapters of the dissertation. Our attention first focuses on a building block of integration theory: trade flow effects among countries party to the agreement.

Chapter 2: What Are the Most Important Partners of the Most Recently Admitted EU Countries?

An extract of this Chapter has been published in:

Economics of Transition (2015), Volume 23, Issue 1, pages 247–292.

In this chapter, we investigate whether the fifth EU enlargement truly generated a trade integration effect for new member states. To answer this question, we construct a database by compiling information regarding trade flows among EU-10 countries and 180 commercial partners in 6 different sectors during 1999–2011. This large body of commercial partners allows us to control for multilateral resistance in the spirit of Anderson and van Wincoop (2003), namely in terms of characteristics of third-party countries. Using the standard gravity model and estimating a difference-in-differences specification, we analyze how joining the European Union has affected the intensity and direction of EU-10 trade flows, not only toward EU-15 countries, but also among both EU-10 and non-EU countries. Though our results show that trade flows among EU-15 and EU-10 countries increased after 2004, the effects were far stronger within the EU-10 group, particularly among EU-10 countries. Briefly, this finding confirms that the historical background of EU-10 countries conditioned their trade integration with EU-15 countries, though not as much as expected, especially in sectors with more technological content.

As previously mentioned, an important step of completing the EU integration process has been implementing the free movement of capital and people, which has notable implications for economic agents, particularly companies and consumers (Baldwin and Wyplosz, 2012). In a single market, companies can select from a range of inputs and factors of production, as well as extend their activity to non-domestic markets, which affords the companies greater efficiency. Yet, this dynamic also implies increased competition, for only the most productive firms can survive in the market, largely due to the selection effect. At the same time, the most productive companies also generate positive spatial spillovers in the territory in which they are located, mostly by enhancing agglomeration forces that affect the location choices of firms and citizens alike. In response to this situation, in Chapter 3 we analyze the interplay of the movement of capital and people (i.e., network effects) represented by foreign direct investment (FDI) and migration inflows in Spain.

Chapter 3: Foreign Direct Investment and Immigration Inflows in Spain

This chapter addresses an important aspect of the European integration process: the determinants of locating FDI inflows in Spain. After joining the EEC in 1986, Spain experienced the entry of significant investment in its domestic economy, mostly from other EEC countries. For these and other foreign investors, Spain exhibited important location advantages, mostly associated with low-cost factors of production. However, the country lost its principal attractiveness as an FDI recipient after EU enlargements to the east in 2004 and 2007, since new member states could offer locations that allowed firms to reduce production costs to an even greater degree.

More specifically, in this chapter we provide a quantitative assessment of the importance of agglomeration economies, network effects, and labor market composition in terms of FDI attractiveness. To conduct our analysis, we create a novel database after adapting and merging information from two micro-data sources: one for firms and the other for workers. Spanning 2005–2012, our analysis encompasses both part of a cycle of expansion until 2007 and part of an economic recession, from 2008 onward. Our results highlight that incoming foreign investors privilege the hiring of medium-skilled workers instead of high-skilled ones, as is often found in empirical research of FDI determinants. By contrast, domestic firms privilege the hiring of low-skilled workers, the demand for which during the 2000s made Spain a preferred destination for many low-skilled immigrants, mostly from Latin American and Eastern European countries. Of course, the combination of both factors reveals the existence of structural problems in the Spanish business environment—namely, that foreign investors are attracted by transitory factors that temporally limit their interest in locations. In effect, this condition prevents Spain from enjoying qualified business environment able to attract more long-term FDI.

Having centered our attention on two highly relevant features of the integration process in Europe, we turn our attention to the CAP, one of the oldest and most controversial EU policies. Established in 1962, the CAP has sought to ensure an adequate food supply and increase the income of farmers. Though important reasons justify regulating the agricultural sector (Staab, 2013), the CAP system has nevertheless been criticized since its early years as an inefficient, expensive policy that benefits farmers with the largest businesses only. Over time, the CAP has thus become the subject of several studies and evaluations, typically with controversial results. In Chapter 4, we thus provide an original empirical analysis to objectively quantify how the CAP affects the level of agricultural production in Spain, a major recipient of CAP subsidies.

Chapter 4: Land Specialization in Spain: The Effects of the Common Agricultural Policy

In this chapter, we investigate the extent to which the CAP has affected the level of agricultural production in Spain. To that end, we perform an empirical analysis of a novel database that considers 50 provinces and 25 crops during 1975–2011. In that sense, Spain is an interesting case for two reasons. First, it is among countries that benefit the most from the CAP budget, meaning that the policy is not simply a marginal source of income for Spanish farmers. Second, since Spain entered the EEC when the CAP had already been established, with these data we can track the evaluation of the level of production over time, thereby distinguishing a period during which Spain was not an EU member and one in which it was and therefore enjoyed CAP subsidies. Following Costinot and Donaldson (2012), the pivotal technique of our strategy involves comparing actual output with potential output, the latter of which derives from a optimization problem relying upon the Ricardian idea of opportunity cost. Ultimately, our results identify improved production efficiency after Spain entered the EEC (1986), and especially after the Fischler reform (2003). We therefore provide evidence that the 2003 CAP reform, which dissolved the linkage between subsidies and production, positively affected real production in Spain, possibly due to changes in incentivizing agricultural production, after which farmers became more effective by relying more on market features instead of subsidy requirements.

The results of our analysis thus reveal the victories and frustrations concerning the effectiveness of the European integration process. Overall, we conclude that integration has created remarkable advantages for EU member states, though there is room to improve its effectiveness for the socioeconomic environment of EU firms and citizens. In our conclusions, we articulate insights into this issue and sketch out feasible extensions of our empirical analysis.

Chapter 2

What are the most Important Partners of the most Recently Admitted EU Countries?[1]

2.1 Introduction

The European Union (EU) has experienced one of the most complete integration processes in the world since its creation in 1957. After its fourth enlargement in 1995, the EU was comprised of 15 countries (EU-15). The fifth enlargement in 2004 added ten member countries (henceforth the EU-10).[2] In principle, trade liberalization is expected to reinforce the intensity of trade flows among EU members. In this context, we aim to investigate the evolution of trade integration among EU members since the fifth enlargement in 2004, namely, between the new and old member countries.

Strong former ties between the EU-10 and communist countries made the Eastern enlargement one of the most challenging from an economic viewpoint. A few EU-10 members (Estonia, Latvia, and Lithuania) were part of the Union of Soviet Socialist Republics (USSR), and others (Bulgaria, the former Czechoslovakia, Hungary, Poland, and Romania), with the USSR, founded the Council for Mutual Economic Assistance (CMEA or COMECON) in 1949. As a result, these countries shared a remarkable interdependence until 1991, when the USSR collapsed. Nevertheless, during the 1990s, the EU-10 and some post-Soviet states continued their strong commercial relationships (Bussière et al., 2008). In this respect, we also investigate how the fifth EU en-

[1]This chapter benefits from comments and feedbacks from the participants of the Applied Lunch Seminar at the UAB (Barcelona, 2011), the International Workshop on Recent Issues in European Economic Integration and EU Enlargement (Brussels, 2011), the VIII Jornadas sobre Integración Económica (Valencia, 2011), the XREPP's Doctoral Day (Barcelona, 2012), the XXXVII SAEe (Vigo, 2012), the seminar in the Department of Economics at the University of Liège (Liège, 2013), the IRES internal seminar in international economics (Louvain-la-Neuve, 2013) and the XXVIII Jornadas de Economía Industrial (Segovia, 2013).

[2]EU-10: Cyprus, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia, and Slovenia.

largement affected the nature of trade flows between the EU-10 and countries of the former Soviet Union (henceforth the FSU).

The instrument selected to develop the empirical analysis is the gravity model. This model has been successful in explaining the intensity and direction of trade flows between countries (Baier and Bergstrand, 2007; Combes et al., 2008; Feenstra, 2004). Our theoretical framework is based on the one developed by Anderson and van Wincoop (2003). We add additional features to the original model in order to facilitate data analysis by, for example, distinguishing *home* goods from *foreign* goods and considering the presence of several economic sectors. The main implications of the model, however, are unchanged. The intensity of trade flows from country i to country j depend on some unilateral and bilateral characteristics (that is, exporter's gross domestic product (GDP), importer's GDP, and trade cost), as well as some multilateral characteristics.

To perform our empirical exercise, we construct a panel of data at reporter–partner–sector level from 1999 to 2011 using three data sources. The 180 potential EU-10 commercial partners are divided into four groups: EU-15, EU-10, FSU, and the rest of the world (ROW), which is used as the reference. From the standard gravity model, we include a select group of dummy variables (DVs) that capture the variation of the intensity of trade flows between the EU-10 and the four commercial partners groups over time. These DVs were designed based on the difference-in-differences (DID) strategy, which is usually used to evaluate the causal effect of implementing a specific program or policy in a target group, the EU-10 in this case.

We investigate how the fifth EU enlargement favored a true trade integration between the EU-10 and EU-15 and how it affected commercial relations between the EU-10 and countries of the FSU. To achieve our objective, we examine variations in the intensity and direction of the EU-10's export and import flows. The advantages of referring to both exports and imports are related to the possibility of detecting the factors that could explain the change in the nature of the EU-10's trade flows. While the determinants of export flows are associated with the degree of competition among the local producers, the determinants of import flows rely more on the preferences and demand of the destination countries.

Two different (but complementary) estimation exercises are proposed to analyze the potential effects of the EU-10's trade flows over time. These exercises differ in how the temporal dimension is managed. In the first exercise, we distinguish two periods: one prior to 2004 (that is, from 1999 to 2003) and one 2004 and later (that is, from 2004 to 2011). The goal is to identify, for each commercial partner group, whether or not the variation of intensity of the EU-10's trade flows between the two periods was relatively higher than, equal to, or lower than those for the other groups. We then propose a complementary exercise to check whether or not the results are sensitive to the time period used as the reference. The second exercise analyzes the trend of the intensity of the EU-10's trade flows with each commercial partner group during the period 1999–2011, using 1999 as the reference year. This technique emphasizes the specific moment when the fifth EU enlargement affected the intensity of the EU-10's trade flows.

We find some remarkable results. We discover that the EU-10's export flows to the EU-15 and EU-10 increased more than to the ROW after 2004, while the exports were redirected from the FSU to the ROW during the years prior to the fifth enlargement. At the sectoral level, we find an interesting heterogeneity in the behavior of the EU-10's export flows to the EU-15. After 2004, the EU-10's export flows to the EU-15 increased more than to the ROW in some sectors (chemicals, food and beverages, and manufactured goods classified by material) but less in others (machinery and vehicles, other manufactured articles, and raw materials and energy). This may have been the result partly because of competitiveness issues. That is, in some sectors the EU-10 products were not sufficiently competitive for the EU-15 markets, and markets outside the EU-15 had to be sought. Among imports, we find trade redirection from the ROW to the EU-15, EU-10, and FSU after 2004. Interestingly, the impact was higher among the EU-10; their import flows from the EU-10 were greater than from the other groups in nearly all sectors. This result could be explained by a strong bias in the demand of EU-10 consumers toward EU-10 products.

Finally, we implement two extensions to prove the robustness of the main results. In the former extension, we deal with the missing values problem. Instead of excluding missing values from the sample, they are replaced by zeros (see, for instance, Gleditsch, 2002). Then, an alternative method, the fixed-effects (FE) Poisson maximum-likelihood (ML) estimator, is implemented in order to handle with a sample where the dependent variable has a large proportion of zero values. The latter extension is associated with the strong relationship between trade flows and foreign direct investment (FDI) discussed in the economic literature (Markusen, 2002 is the main reference). We determine whether, after controlling for bilateral FDI, the coefficients of interest change their magnitude and statistical significance. Both extensions confirm that our results are robust.

This chapter is organized as follows. Section 2.2 is a review of the relevant literature. Section 2.3 is an overview of the international relationships between the EU and other countries, especially those of the FSU. Section 2.4 presents the theoretical framework. Section 2.5 focuses on the empirical strategy, the description of the data and the explanation of the two different (but complementary) estimation exercises. Section 2.6 presents and discusses the results obtained from the two estimation exercises, while Section 2.7 presents the extensions, proving the robustness of the main results. Finally, Section 2.8 concludes and addresses policy recommendations.

2.2 Literature Review

The Eastern enlargement was not an unpredictable event; since the beginning of the 1990s, some Central and Eastern European countries (CEECs)[3] expressed interest in joining the EU. For this reason, a relevant part of the literature discusses the degree of trade integration among Western and Eastern Europe during the 1990s. Many studies confirm that there was an important trade integration process among these European areas during the 1990s (Abraham and Konings, 1999; Bussière et al., 2008; Brenton and Gros, 1997; Fontagné et al., 1999; Gros and Steinherr). Others argue, however, that this integration was far from complete (Faucompret and Vandenbussche 1999; Paas, 2003).

In particular, Gros and Steinherr (1995) explain that during the 1980s, the trade activity of Central and Eastern Europe (CEE) occurred mostly within the COMECON bloc and that trade among socialist economies accounted for approximately 60 to 70 percent of all CEE trade. From 1989 to 1992, however, CEE trade was redirected toward developed countries, which by 1992 accounted for two thirds of all CEE trade. Brenton and Gros (1997) study the transition processes of the CEECs and post-Soviet states after the dissolution of the USSR by analyzing trade in the intra-Organisation for Economic Co-operation and Development. They conclude that, in terms of geographical composition, the trade of some CEECs (the Czech Republic, Hungary, and Poland) was indistinguishable from a few EU-15 members (Austria and Spain).[4] They thus confirm that in some CEECs there was trade reorientation from the former COMECON partners toward the Western countries, particularly the ones belonging to the EU. Bussière et al. (2008) show that Russia and Ukraine remained important trading partners for CEECs at the end of the 1990s. Conversely, Faucompret and Vandenbussche (1999) consider that the trade reorientation during the 1990s was incomplete because the EU followed overly restrictive measures for goods imported from non-EU countries. In fact, the EU shifted from being a net importer to a net exporter between 1990 and 1996, meaning that the intensity of the EU's export flows to the

[3]Bulgaria, the former Czechoslovakia, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, and Slovenia.

[4]Gros and Gonciarz (1996) already showed that the distribution if trade of the CEECs was quite close to those of EU countries of similar size.

CEECs experienced a sharp increase.

Similarly, other studies focus on the effects of European trade agreements on the intensity and direction of trade flows among European countries. Herderschee and Qiao (2007) confirm that the Europe Agreements (EAs) contributed significantly to bilateral trade flows between the EU and some CEECs. Some years later, Egger and Larch (2011) confirm that the EAs fostered trade between the EU and CEECs and reduced trade flows between CEECs and other commercial partners, namely the post-Soviet states and former Yugoslavia. They also find negative intra-group effects—that is, the intensity of trade flows within the EU and within the CEECs decreased.

Despite the abundance of studies on the effects of the trade liberalization process during the 1990s, those that focus on the potential effects on the intensity and direction of the EU-10's trade flows after 2004 are limited (Hornok, 2010 and Antimiani and Costantini, 2013 are two of them). In light of the previous discussion, the main contribution of this chapter is to provide empirical evidence of the potential effects of the fifth EU enlargement on the EU-10's trade flows.

From a technical perspective, this study acknowledges the gravity framework. In the standard gravity model (Tinbergen, 1962), bilateral trade flows are positively correlated with the size of each partner and negatively affected by trade cost. The size of the countries is often measured by the GDP, while trade cost is measured by the distance between countries. This model has been characterized by the quality of its empirical results. It has been used to estimate the impact of common borders (Chen, 2004; McCallum, 1995; Nitsch, 2000), preferential trading blocs (Baier and Bergstrand, 2007; Carrère, 2006), and currency union (Glick and Rose, 2002; Rose and van Wincoop, 2001), among other things, on the intensity and direction of trade flows.

One of the major weakness of gravity models was the absence of any theoretical foundation endorsing the setting. Then, economic literature counted on several contributions that filled up that gap (see, for instance, Anderson, 1979; Anderson and van Wincoop, 2003; Baier and Bergstrand, 2007; Bergstrand, 1985; Chaney, 2008; Chaney, 2013; and Morales et al., 2014). Anderson (1979) was the first in developing a theoretical foundation for the gravity model in international trade. He assumes that preferences are identical and homothetic among regions and that products are differentiated by place of origin, which is known as Armington assumption. Then, each region is completely specialized in the production of its own good (there is a good for each country). Through the manipulates of an expenditure function, he obtain interesting predictions that detects that the usual estimator of the gravity equation may be biased. Based on the Anderson (1979)'s foundations, Anderson and van Wincoop (2003) provide one of the most complete frameworks for the gravity model in both directions, theoretical and empirical. Their model predicts that trade flows between two regions depend not only on the trade costs between these two regions but also on the trade costs between these two regions and the ROW, which became known as the *multilateral* resistance term. They prove that the standard gravity model, which considers only bilateral trade costs, produces biased results and, therefore, yields misleading interpretations. In this study we extend their model distinguishing *home* goods from *foreign* goods and by considering several economic sectors.

Implementing a gravity framework to study the dynamics resulting from preferential trade agreements (PTAs) is not new. Some studies (for example, Frankel, 1997, chapter 5) introduce two DVs to the standard model equation to capture information about both trade creation and trade diversion. Carrère (2006), Soloaga and Winters (2001), and Westerlund and Wilhelmsson (2011) introduce a third DV to account for the diversion of exports. Unlike these studies, our sample does not allow us to identify the effects of trade creation or trade diversion because our reporter countries are limited to those belonging to the EU-10.[5] Taking into account this limitation, we devise an econometric strategy that exploits the existence of several EU-10's commercial partner groups (EU-15, EU-10, FSU, and ROW) and, at temporal dimension, recognizes the date of their acceptance into the EU (2004). Then, implementing a technique of policy evaluation, the DID strategy, we are able to identify if, after inclusion, they experienced an important variation in the intensity and direction of trade flows.

On the technical side, economists have attempted to improve the specifications of the empirical gravity model to fit theoretical advances. According to Baltagi (2008), the characteristics of the panel econometric framework reduce the probability of obtaining biased results. The first gravitational studies using longitudinal data appeared in the 1990s. Mátyás (1997), for example, estimates the volume of exports in eleven countries of the Asia-Pacific Economic Cooperation from 1982 to 1994. He selects two models, one that does not account for fixed effects and one that includes exporter, importer, and year fixed effects. The proposed exercise allows us to detect important differences in the magnitude of the estimated coefficients of the explanatory variables (GDP, population, foreign currency reserves, and real exchange rates) in the two models. Some years later, Egger and Pfaffermayr (2003) exploits Mátyás' data and gravity equation to prove that the correct specification of the gravity model should also include bilateral interaction effects, or exporter–importer fixed effects. Though this type of fixed effects does not allow users

[5]To correctly interpret Vinarian trade creation and trade diversion, we needed a sample with countries belonging to some PTA and, for a control, countries not belonging to any trading bloc.

to evaluate time-invariant variables such as distance, border, and language, it does capture all unobservable heterogeneity. Baltagi et al. (2003) also include country-time DVs to control for trends specific to each country. Baldwin and Taglioni (2006) discuss that both the country-pair and the country-time DVs can eliminate what they call 'gold medal' bias (the omitted variable bias or the multilateral resistance term). Nevertheless, they also remark that the inclusion of all these fixed effects generate many DVs and, consequently, an important loss in degrees of freedom. In this study, we introduce different types of fixed effects in our econometric specification. The sectoral dimension of the database (ten sections of the Standard International Trade Classification (SITC)) will be relevant in describing and interpreting parts of our results.

Finally, another important concern in the empirical trade literature is the presence of zero values, which occur frequently when considering bilateral trade flows in a sample with a large number of countries. Santos Silva and Tenreyro (2006) demonstrate that the standard log-linearized gravity equation estimated by ordinary least squares (OLS) produces biased results because of the heteroskedasticity problem and truncation of the data (the log-linearization model drops the zero values). They propose a Poisson pseudo-ML estimation technique to solve both issues. Similarly, Westerlund and Wilhelmsson (2011) show that the standard gravity model estimated by OLS produces biased and inefficient results in the presence of heteroskedasticity and zero trade flow values. They propose a FE Poisson ML estimator using a panel of countries to analyze the effects of the adhesion of Austria, Finland, and Sweden to the EU in 1995 via the intensity of import flows. In light of these results, we perform a robustness exercise, adhering to these methods, to control for the truncation problem in our database.

2.3 International Context: An Overview

In this section, we provide a synthetic overview of the historical background in Europe, starting with the dissolution of the USSR (1991) and ending in the early 2000s. We first discuss relations between the EU and some CEECs and then describe relations between the EU and countries of the FSU, especially the Russian Federation, the leading member of the USSR.

The CEECs were accustomed to having strong political and economic ties with the USSR. In 1949 the USSR, Bulgaria, Czechoslovakia, Hungary, Poland, and Romania founded the COMECON, with the primary goal of establishing strong economic relationships between socialist countries. Among the EU-10, only the Baltic countries (Estonia, Latvia, and Lithuania) were part of the USSR, and they became independent states in 1991. Ten countries in the CEECs applied for EU membership in the early 1990s. As a step toward complete integration, the EU decided to sign EAs with them to progressively establish bilateral free trade for manufactured products and ultimately remove all trade barriers between the EU and these CEECs by the end of the 1990s.[6] In addition, the EAs also aimed to shift national economies from planned to market economies. Nevertheless, according to Baldwin and Wyplosz (2012), these free trade agreements were incomplete because the EU maintained tariffs and trade restrictions on some industrial products.[7] Importantly, in June 1993, the European Council advanced the integration process by deciding that countries signing the EAs could become official members once they fulfilled the Copenhagen criteria. In May 2004, the fifth EU enlargement occurred and ten new countries received membership.

Since the end of the 1990s, the EU concluded similar partnership and cooperation agreements (PCAs) with Russia and nine Newly Independent States.[8] These agreements aimed to promote trade and investment among both EU and FSU countries.[9] At the St. Petersburg Summit in May 2003, the EU and Russia agreed to reinforce their cooperation by gradually creating four common spaces in the framework of the PCA: a common economic space; a common space of freedom, security, and justice; a space of cooperation for external security; and a space of research, education, and cultural exchange. Furthermore, the EU was a strong supporter of Russia joining the World Trade Organization from the beginning of the process until its accession in 2012.

Despite past ties among CEECs and Russia, some controversial situations occurred during the 2000s. Russia imposed a ban on Polish meat imports from 2005 to 2007 because of allegations that Poland exported low-quality, unsanitary meat products. As a consequence, Poland blocked a proposed bilateral treaty between the EU and Russia at the Samara Summit in May 2007. The Baltic countries also reacted against these talks: Estonia complained of Russian cyber attacks upon its government, news media, and banking websites; Lithuania faced a Russian oil blockage; and Latvia opposed Russia's Baltic pipeline plan on environmental grounds (Rettman, 2007).

[6] The EU signed EAs in 1991 (Poland and Hungary), 1993 (Bulgaria, the Czech Republic, Romania, and Slovakia), 1995 (Estonia, Latvia, and Lithuania), and 1996 (Slovenia). The agreements came into force approximately two or three years afterward (EC, 2001a).

[7] A group of 'sensitive' products (textiles, coal and steel products, and agricultural products) continued to receive strong protection.

[8] The Republic of Armenia, the Republic of Azerbaijan, Georgia, the Republic of Kazakhstan, the Kyrgyz Republic, the Republic of Moldova, Ukraine, the Republic of Uzbekistan, and Tajikistan.

[9] Source: http://europa.eu/legislation_summaries/external_relations/index_en.htm.

2.4 Theoretical Framework

This section aims to sketch a theoretical setting to endorse the interpretation of our empirical exercise. The building blocks of our gravity model owe much to the one presented by Anderson and van Wincoop (2003). Their model relies on three key assumptions: each region is specialized in the production of only one (tradable) good, preferences are identical and homothetic, and the trade costs are symmetric. We adopt the second and third assumptions, but we add some extensions, namely the distinction between home and foreign goods and the existence of several economic sectors. As Anderson and van Wincoop (2003)'s setting, we only consider the demand side of any given economy. We assume that countries are endowed with a fixed supply of goods, which depends on the total capacity of the economy, or total income.[10]

Preferences We consider C symmetric countries and, in each country, we distinguish two types of goods: the home good (H) and the foreign good (F). Domestic and foreign goods are not perfect substitutes: there exists a home bias in the consumption decisions. Preferences, identical and homothetic across countries, are represented by the following Cobb-Douglas (CD) utility function:

$$U_j = H_j^{1-\mu} F_j^{\mu} \qquad 0 < \mu < 1 \quad , \tag{2.1}$$

Then, the utility of the representative consumer in country j depends on the quantity of the home good (H_j), the quantity of the foreign good or the total imports (F_j), and the elasticities of the home good and the foreign good ($1-\mu$ and μ, respectively). At aggregate level, the following condition has to be fulfilled:

$$H_j P_{H_j} + F_j P_{F_j} \leq Y_j \quad \forall j \quad , \tag{2.2}$$

where P_{H_j} is the price index of the home good, P_{F_j} is the price index of the foreign good, and Y_j is the country j's income. The previous constraint states that the consumers in country j cannot spend more than the country j's income.

The home good (H_j) is consider as the *numéraire* while the foreign good (F_j) is represented by a CD-type function:

$$F_j = \prod_{s=1}^{S} \left(f_j^s\right)^{\gamma^s} = \left(f_j^1\right)^{\gamma^1} \times \left(f_j^2\right)^{\gamma^2} \times ... \times \left(f_j^S\right)^{\gamma^S} , \qquad s = 1, ..., S \ \text{ and } \sum_{s=1}^{S} \gamma^s = 1 \quad , \tag{2.3}$$

[10] We then assume that the representative firm is efficient, meaning always produces at its maximum capacity.

where f_j^s is the quantity imported by country j in the sector s and γ^s is the elasticity for each imported sector s.[11] According to the Armington (1969)'s assumption, products are differentiated by place of origin. The idea is that for each sector s, each country produces a representative commodity. Therefore, f_j^s is a composite good composed by a group of representative commodities, namely one for each trading country.[12] Along this section we refer f_j^s as the (foreign) composite good s, and it takes the form of a constant elasticity of substitution (CES) function:

$$f_j^s = \left[\sum_{i=1, i \neq j}^{C-1} \left(\beta_{ij}^s c_{ij}^s \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \ , \ \ i = 1, ..., j, ..., C-1 \ , \quad s = 1, ..., S \ , \ \text{ and } \ 1 < \sigma < \infty \ , \tag{2.4}$$

where c_{ij}^s is the consumption in country j of the representative commodity produced by country i in sector s, β_{ij}^s is the weight attributed by consumers in country j to the representative commodity produced by country j in sector s, and σ is the elasticity of substitution among commodities.[13]

Aggregate Demand In order to obtain the country j's demand function for the home and foreign goods, country j consumers maximize the utility (2.1) subject to the budget constraint (2.2). So we get the following maximization problem:

$$\begin{aligned} \max U_j &= H_j^{1-\mu} F_j^{\mu} \\ \text{s.t. } Y_j &= H_j P_{H_j} + F_j P_{F_j} \ . \end{aligned} \tag{2.5}$$

We build the Lagrangian multiplier ($\mathcal{L}_j^1$) and solve for first order conditions (FOCs).

$$\mathcal{L}_j^1 = H_j^{1-\mu} F_j^{\mu} + \lambda (Y_j - H_j P_{H_j} - F_j P_{F_j}) \tag{2.6}$$

FOCs

$$\frac{\partial \mathcal{L}_j^1}{\partial H_j} = 0 \rightarrow (1-\mu) H_j^{-\mu} F_j^{\mu} - \lambda = 0 \tag{2.6.1}$$

$$\frac{\partial \mathcal{L}_j^1}{\partial F_j} = 0 \rightarrow \mu H_j^{1-\mu} F_j^{\mu-1} - \lambda P_{F_j} = 0 \tag{2.6.2}$$

[11] We assume that sectors are complementary in the sense that the representative consumer has to consume all these goods ($f^s \neq 0 \ \forall s$).

[12] For each sector s, country j's consumers import $C-1$ (representative) commodities.

[13] The elasticity of substitution indicates how different (or similar) are the representative commodities. A low value indicates that commodities are not good substitutes while a high value indicates that commodities can be easily substituted. The extreme case, $\sigma \rightarrow \infty$, indicates that commodities are perfect substitutes.

$$\frac{\partial \mathcal{L}_j^1}{\partial \lambda} = 0 \rightarrow Y_j - H_j P_{H_j} - F_j P_{F_j} = 0 \tag{2.6.3}$$

After some algebraic manipulation between equations (2.6.1) and (2.6.2), we obtain the following relationship between the home good and foreign good:

$$\frac{F_j}{H_j} = \frac{\mu}{1-\mu} \frac{P_{H_j}}{P_{F_j}} \quad . \tag{2.6.4}$$

From equation (2.6.4) we obtain:

$$\text{(a) } H_j P_{H_j} = \frac{1-\mu}{\mu} F_j P_{F_j} \text{ and (b) } F_j P_{F_j} = \frac{1-\mu}{\mu} H_j P_{H_j} \quad . \tag{2.6.4'}$$

Plugging equation (2.6.4'a) into equation (2.6.3), we obtain that the demand function for the *foreign* good is equal to

$$F_j^* = \frac{\mu Y_j}{P_{F_j}} \quad . \tag{2.6.5}$$

As one can expect, the demand function of the foreign good depends positively on the income and negatively on its price index. Moreover, the share of expenditure on the foreign good is equal to μ, namely the power of the foreign good in the utility function (2.1).[14]

Finally, plugging equation (2.6.4'b) into (2.6.3), we obtain that the demand function for the *home* good s equal to

$$H_j^* = \frac{(1-\mu) Y_j}{P_{H_j}} \quad , \tag{2.6.6}$$

which depends positively on the income devoted on the home good ($(1-\mu) Y_j$) and negatively on its price index.

Prices and Transport Costs Countries trade because their own citizens love varieties: they like to consume a mix of commodities. For any given representative commodity belonging to the sector s and produced by the country i (namely c_{ij}^s), let p_{ij}^s the price that country j consumers have to pay and let p_i^s the price at home before delivery.[15] Following the idea of iceberg trade costs (Samuelson, 1954), trade costs are modelled as follows:

$$p_{ij}^s = \tau_{ij} p_i^s \ , \ \tau_{ij} > 1 \text{ and } i \neq j \quad , \tag{2.7}$$

[14] We assume that this share is independent of the country's characteristics (income or income per capita).

[15] The price p_{ij}^s includes transport costs from country i to country j on a c.i.f. basis (that is, considering cost, insurance and freight), while the price p_i^s is net of any transport cost on a f.o.b. basis (that is, free of board).

where τ_{ij} is a positive parameter greater than one. The idea behind the iceberg trade costs is that when a good is traded form country i to country j, the amount $(\tau_{ij} - 1)$ melts down along the way and the consumers of the importing country pay for it. Therefore, the nominal value of exports in sector s from country i to country j (that is, country j's payments to country i) is equal to the sum of the value of production at origin $(p_i^s c_{ij}^s)$ and the trade costs $((\tau_{ij} - 1)p_i^s c_{ij}^s)$,

$$x_{ij}^s = p_i^s c_{ij}^s + (\tau_{ij} - 1)p_i^s c_{ij}^s = \tau_{ij} p_i^s c_{ij}^s = p_{ij}^s c_{ij}^s \quad . \tag{2.8}$$

Demand for the Representative Commodity In order to obtain the country j's demand function for each representative commodity produced by country i in sector s (namely c_{ij}^s), we consider the following maximization problem for each (foreign) composite good s:

$$\begin{aligned} \max f_j^s &= \left(\sum_{i=1; i \neq j}^{C-1} \left(\beta_{ij}^s c_{ij}^s \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \\ \text{s.t.} \;\; (\mu \cdot \gamma^s) Y_j &= \sum_{i=1; i \neq j}^{C-1} (c_{ij}^s p_{ij}^s) \qquad , \end{aligned} \tag{2.9}$$

where the representative country j's consumer maximize equation (2.4) subject to the budget constraint for the (foreign) composite commodity s.[16] We build the Langrangian multiplier $(\mathcal{L}_j^2)$ and solve for FOCs.

$$\mathcal{L}_j^2 = \left(\sum_{i=1; i \neq j}^{C-1} \left(\beta_{ij}^s c_{ij}^s \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} + \lambda \left((\mu \cdot \gamma^s) Y_j - \sum_{i=1; i \neq j}^{C-1} (c_{ij}^s p_{ij}^s) \right) \tag{2.10}$$

FOCs

$$\frac{\partial \mathcal{L}_j^2}{\partial c_{ij}^s} (i \neq j) = 0 \rightarrow \left(f_j^s \right)^{\frac{1}{\sigma}} \left(\beta_{ij}^s \right)^{\frac{\sigma-1}{\sigma}} (c_{ij}^s)^{-\frac{1}{\sigma}} - \lambda p_{ij}^s = 0 \tag{2.10.1}$$

$$\frac{\partial \mathcal{L}_j^2}{\partial \lambda} = 0 \rightarrow (\mu \cdot \gamma^s) Y_j - \sum_{i=1; i \neq j}^{C-1} \left(c_{ij}^s p_{ij}^s \right) = 0 \tag{2.10.2}$$

$$\text{where } f_j^s = \left(\sum_{i=1; i \neq j}^{C-1} \left(\beta_{ij}^s c_{ij}^s \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

Rearranging (2.10.1) we obtain the following equation:

[16] The term $\mu \cdot \gamma^s$ is the income share devoted to each (foreign) composite good s.

$$c_{ij}^s = (\lambda)^{-\sigma} f_j^s \left(\beta_{ij}^s\right)^{\sigma-1} (p_{ij}^s)^{-\sigma} \quad , \tag{2.10.1'}$$

and multiplying both sides by the price p_{ij}^s we obtain:

$$c_{ij}^s p_{ij}^s = (\lambda)^{-\sigma} f_j^s \left(\frac{p_{ij}^s}{\beta_{ij}^s}\right)^{1-\sigma} \quad . \tag{2.10.1''}$$

Plugging equation (2.10.1") into (2.10.2), we obtain the following expression for $(\lambda)^{-\sigma} f_j^s$:

$$(\lambda)^{-\sigma} f_j^s = \frac{(\mu \cdot \gamma^s) Y_j}{\sum_{i=1;i\neq j}^{C-1} \left(\left(\beta_{ij}^s\right)^{\sigma-1} (p_{ij}^s)^{1-\sigma}\right)} \quad . \tag{2.10.2'}$$

Finally, plugging equation (2.10.2') into equation (2.10.1'), we obtain that the country j's demand function for each representative commodity produced in country i in sector s is equal to

$$c_{ij}^s = \frac{\left(\beta_{ij}^s\right)^{\sigma-1} (p_{ij}^s)^{-\sigma}}{\sum_{i=1;i\neq j}^{C-1} \left(\left(\beta_{ij}^s\right)^{\sigma-1} (p_{ij}^s)^{1-\sigma}\right)} (\mu \cdot \gamma^s) Y_j \quad . \tag{2.11}$$

The country j's demand of the each representative commodity is a function that depends on the price (p_{ij}^s), the weight attributed by consumers in country j to the representative commodity produced by country i in sector s (β_{ij}^s), the level of expenditure on the (foreign) composite good s ($(\mu \cdot \gamma^s) Y_j$), and the scaled prices of all the commodities belonging to the (foreign) composite good s (this is the denominator).

Price Index of (Foreign) Composite Good s We can compute the price index of the (foreign) composite good s replacing equation (2.11) into the CES-type index (2.4) and arranging terms.

$$f_j^s = \left[\sum_{i=1;i\neq j}^{C-1} \left[\beta_{ij} \left(\frac{\left(\beta_{ij}^s\right)^{\sigma-1} (p_{ij}^s)^{-\sigma}}{\sum_{i=1;i\neq j}^{C-1} \left(\left(\beta_{ij}^s\right)^{\sigma-1} (p_{ij}^s)^{1-\sigma}\right)} (\mu \cdot \gamma^s) Y_j\right)\right]^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \quad \rightarrow$$

$$\rightarrow f_j^s = \left[\frac{\sum_{i=1;i\neq j}^{C-1} \left(\left(\beta_{ij}^s\right)^{\sigma-1} (p_{ij}^s)^{1-\sigma} \right)}{\left[\sum_{i=1;i\neq j}^{C-1} \left(\left(\beta_{ij}^s\right)^{\sigma-1} (p_{ij}^s)^{1-\sigma} \right) \right]^{\frac{\sigma-1}{\sigma}}} \right]^{\frac{\sigma}{\sigma-1}} (\mu \cdot \gamma^s)\, Y_j \rightarrow$$

$$\rightarrow f_j^s = \left(\sum_{i=1;i\neq j}^{C-1} \left(\left(\beta_{ij}^s\right)^{\sigma-1} (p_{ij}^s)^{1-\sigma} \right) \right)^{\frac{1}{\sigma-1}} (\mu \cdot \gamma^s)\, Y_j \rightarrow$$

$$\rightarrow (\mu \cdot \gamma^s)\, Y_j = f_j^s \left(\sum_{i=1;i\neq j}^{C-1} \left(\left(\beta_{ij}^s\right)^{\sigma-1} (p_{ij}^s)^{1-\sigma} \right) \right)^{\frac{1}{1-\sigma}} \tag{2.12}$$

This last equation states that the expenditure on the (foreign) composite commodity s must be equal to the amount imported and consumed times its price index. So we can define the price index of the (foreign) composite good s as

$$P_{f_j^s} = \left(\sum_{i=1;i\neq j}^{C-1} \left(\left(\beta_{ij}^s\right)^{\sigma-1} (p_{ij}^s)^{1-\sigma} \right) \right)^{\frac{1}{1-\sigma}} \quad . \tag{2.13}$$

Taking into account the price index and the transport costs, the country j's demand for any representative commodity produced by country i in sector s (c_{ij}^s), can be expressed as follows:

$$x_{ij}^s = \left(\frac{\tau_{ij} p_i^s}{\beta_{ij}^s P_{f_j^s}} \right)^{1-\sigma} (\mu \cdot \gamma^s)\, Y_j \quad . \tag{2.11'}$$

Market Clearing Condition The equilibrium of the model relies on the market clearance of the external sector (namely, trade balance for each sector s):

$$(\mu \cdot \gamma^s)\, Y_i = \sum_{j=1;j\neq i}^{C-1} x_{ij}^s = \sum_{j=1;j\neq i}^{C-1} \left(\frac{\tau_{ij} p_i^s}{\beta_{ij}^s P_{f_j^s}} \right)^{1-\sigma} (\mu \cdot \gamma^s)\, Y_j \quad . \tag{2.14}$$

This condition implies that the value of the country i's import flows (left-side) must be equal to the value if the country i's export flows.[17]

From equation (2.14) we solve for the term $(p_i^s)^{1-\sigma}$:

[17] The idea of this assumption is that in the long-run the external sector must be in equilibrium (value of exports equal to value of imports).

$$(p_i^s)^{1-\sigma} = \frac{(\mu \cdot \gamma^s)\, Y_i}{\sum\limits_{j=1;j\neq i}^{C-1} \left[\left(\frac{\tau_{ij}}{\beta_{ij}^s P_{f_j^s}} \right)^{1-\sigma} (\mu \cdot \gamma^s)\, Y_j \right]} \quad . \tag{2.15}$$

Finally, plugging equation (2.15) into demand equation (2.11'), we obtain the following gravity model:

$$x_{ij}^s = (\mu \cdot \gamma^s)^2\, Y_i Y_j \left(\frac{\tau_{ij}}{\beta_{ij}^s P_{f_j^s}} \right)^{1-\sigma} \Omega_i \quad , \tag{2.16}$$

$$\text{where } (\Omega_i) = \frac{1}{\sum\limits_{j=1;j\neq i}^{C-1} \left[\left(\frac{\tau_{ij}}{\beta_{ij}^s P_{f_j^s}} \right)^{1-\sigma} (\mu \cdot \gamma^s)\, Y_j \right]} \quad .$$

The exports from country i to country j in sector s (that is, the country j's imports of the representative commodity produced by country i in sector s) depend on the share of expenditure devoted to the (foreign) composite good s $(\mu \cdot \gamma^s)$, the exporter's income (Y_i), the importer's income (Y_j), the transport costs between country i and country j (τ_{ij}), the weight attributed by consumers in country j to the representative commodity produced by country i in sector s (β_{ij}^s), the price index of the (foreign) composite good s $(P_{f_j^s})$, and the elasticity of substitution between products (σ).[18]

In the spirit of Anderson (1979) and Anderson and van Wincoop (2003), the characteristics of third countries are important to explain the intensity of trade flows among any two commercial partners. The term $Y_j \tau_{ij}^{1-\sigma}$ describes how attractive is market j for country i, that is, how important is the country j's demand (after considering the trade costs). The model also incorporates a multilateral variable (Ω_i), a term which captures the characteristics of third countries. This means that if, on average, ROW countries are more attractive than the importing country j, the export flows from country i *to* country j will be negatively affected. This is closely related to the *multilateral* resistance term introduced by Anderson and van Wincoop (2003). In the empirical analysis, we use reporter–partner–sector fixed effects (when all the sectors are estimated together) and reporter-partner fixed effects (when each sector is estimated separately) in order to control for the invariant-unobservable characteristics.

[18]The impact of the elasticity of substitution is the following: $\sigma > 1$ implies that $(1 - \sigma) < 1$, so the export flows from country i to country j in sector s depend positively on β_{ij}^s and $P_{f_j^s}$ and negatively on τ_{ij}. In addition, export flows depend negatively on σ because a greater elasticity of substitution means that the representative commodity produced by country i in sector s (namely c_{ij}^s) is less differentiated.

2.5 Empirical Strategy

In this section we present the empirical strategy implemented to assess the potential effects of the fifth EU enlargement on the EU-10's trade flows. We present the empirical strategy in two parts. First, we describe the data and discuss empirical evidence of the evolution of the intensity of the EU-10's trade flows with respect to the four commercial partner groups. Second, we present the two different (but complementary) estimation exercises.

2.5.1 Data

To implement our empirical strategy we use trade flow data (exports and imports) between the EU-10 countries and 180 countries (25 EU countries and 155 non-EU countries)[19] during the period 1999–2011. For each country-pair entry we distinguish between the reporter and the partner. The reporter is one of the ten countries composing the EU-10 (reporting the total value of exports and imports), and the partner is any of the 180 potential commercial partners.[20]

The database has been built using three sources of data. Information on trade flows is extracted from Eurostat database *EU Trade since 1988 by SITC*.[21] The SITC divides exports and imports into ten broad sections, numbered 0 through 9. Following the sectoral classification implemented in Eurostat (2012), the sections are aggregated into six groups: chemicals, food and beverages (F&B), machinery and vehicles, manufactured goods classified by materials, other manufactured articles, and raw materials and energy.[22] Nominal GDP data are taken from the *World Economic Outlook* database published by the International Monetary Fund (IMF, 2012). To analyze trade data and GDP in the same nominal currency (current euro), we use the bilateral exchange rate (EUR/USD) reported by Eurostat. Finally, geographical variables for the gravity model (distance and the dummy variable indicating contiguity) are extracted from the CEPII database *GeoDist* (Mayer and Zignago, 2011). Using this information, we construct a panel of data at the reporter-partner-sector level for the period 1999–2011.[23] The panel structure of the database allows us to better control for heterogeneity because the trade flow intensity between

[19] The list of countries appears in Appendix 2.9 (Tables 2.12 and 2.13)

[20] In order to take into consideration the theoretical condition of symmetry among countries, our sample of reporter countries include only the new member states joining EU in the fifth enlargement (2004).

[21] This database distinguishes between zeros and missing values. When a country-pair trade flow is equal to zero, this means 'less than half the final digit shown and greater than real zero;' that is, higher than 0 and less than 0.5 euros. But when a value is missing, this means that the information is unavailable.

[22] Aggregation is as follows: SITC sections 0 and 1 comprise F&B, 2 and 3 comprise raw materials and energy, 5 comprises chemicals, 6 comprises manufactured goods classified by materials, 7 comprises machinery and vehicles, and 8 comprises other manufactured articles.

[23] This period was selected for two reasons. First, Eurostat provides information about the EU-10's trade flows only since 1999. Second, the EU-10 countries went through a transition process during the 1990s. Our objective is to study the impact of becoming a member of the EU rather than the transition process.

any given two countries in any specific economic activity could be followed through time. The full sample includes 139,620 potential observations. After dropping missing values as in Harris et al. (2012), the sample is reduced to 87,125 entries in the case of the exports and 79,848 entries in the case of the imports.[24]

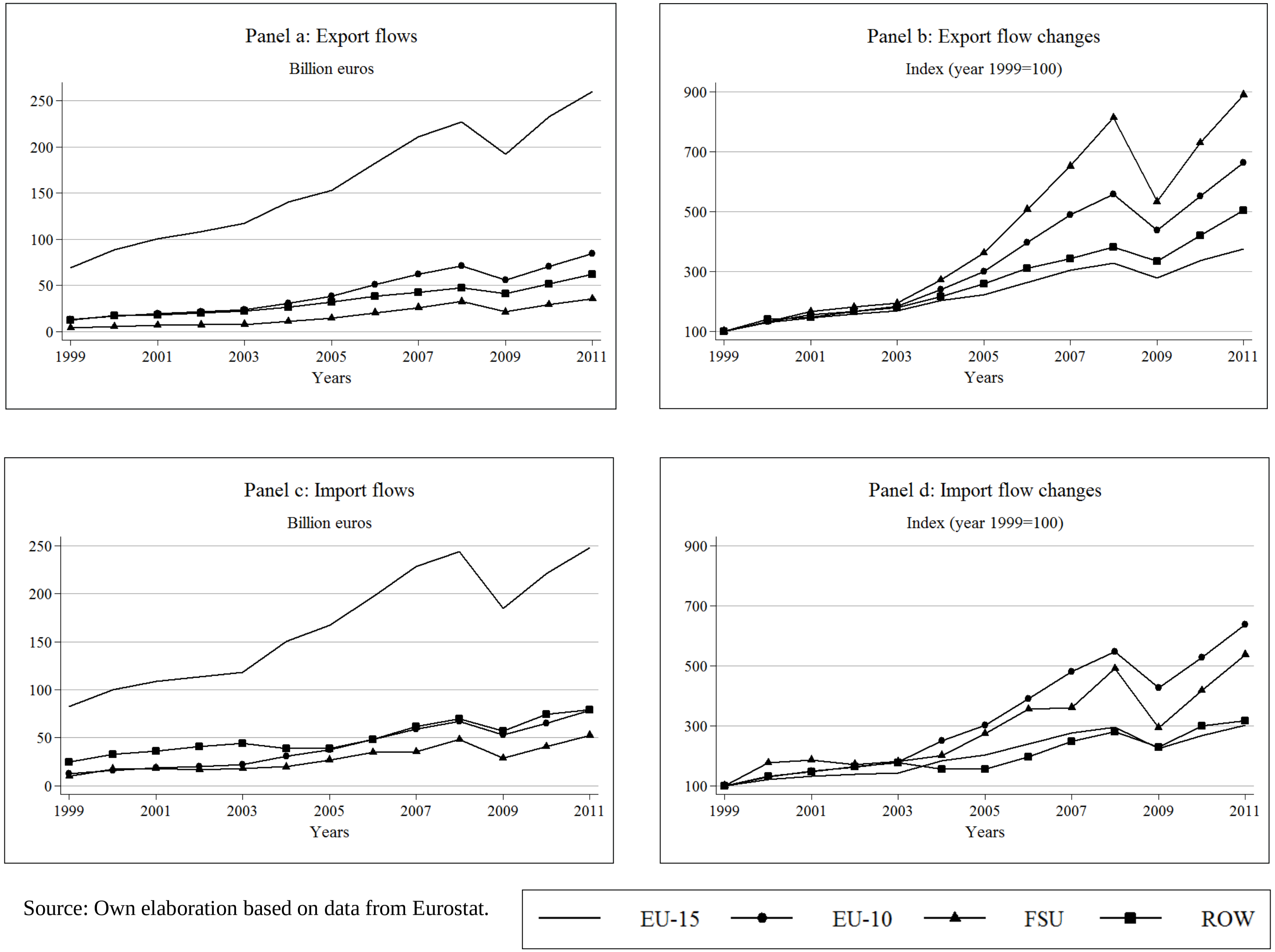
Some empirical evidence of the evolution of the intensity of the EU-10's trade flows is presented in Figure 2.1 (exports in upper panels and imports in bottom panels). Panel a is a plot of the EU-10's export flows (in billions of euros) for the period 1999–2011. During the entire period, exports to the EU-15 were higher than to other groups. The scale of the y-axis does not allow proper appreciation of the trends in exports to the FSU and ROW. Panel b presents a correction for this scale effect, normalizing all values so that the exports from the EU-10 to any group is equal to 100 in 1999. With this correction, we can see that since the fifth EU enlargement in 2004, export growth rates considerably diverged among the groups. Exports to the EU-15 grew at the slowest rate. This could be explained by the EAs signed between the EU-15 and some CEECs during the 1990s, which might have anticipated the trade effect of the Eastern enlargement a few years before the official entry. This could also be because the initial level of the EU-10's export flows to the EU-15 was higher than to the other destinations, influencing export growth rates throughout the period. Panel 1b is replicated by sector (see Appendix 2.9, Figure 2.4). Exports to the FSU experienced the highest growth rate in three sectors (chemicals, machinery and vehicles, and manufactured goods classified by materials). The EU-10's greatest export growth rate in the F&B sector was to the EU (EU-15 and EU-10). Exports to the ROW experienced the highest growth rate in the raw materials and energy sector.[25]

We use the same exercise on the import data. Panel c is a plot of the EU-10's import flows (in billions of euros) for the period 1999–2011. Imports from the EU-15 were higher than from other groups. To avoid the scale effect, the data were normalized in the same way as the export data, and are presented in panel d. Import growth rates varied among commercial partner groups at the beginning of the period and intensified after 2004. The intensity of the EU-10's import flows from the EU-10 experienced the highest growth rate. This result, generalized for all sectors, was specially notable in machinery and vehicles and the other manufactured articles sectors (see Appendix 2.9, Figure 2.5).

[24]Missing values indicate that data are unavailable for either trade flows or GDP.

[25]According to the EU official webpage (http://europa.eu/index_en.htm), Poland is rich in natural mineral resources such as iron, zinc, copper, and rock salt, while Hungary is endowed with other natural resources (for example, bauxite, coal, and natural gas).

Figure 2.1: EU-10 trade flows by trading partner



Source: Own elaboration based on data from Eurostat.

2.5.2 Econometric Specification

In this subsection we present the two different (but complementary) estimation exercises implemented to study the potential effects of the fifth EU enlargement on the intensity and direction of the EU-10's trade flows.

Our gravity equation (2.16) identifies that the trade intensity form country i to country j relies on selected unilateral and bilateral features, as well as some *multilateral* effects. Taking natural logarithms of (2.16) we obtain the following equation:

$$\ln X_{ij}^s = 2\ln(\mu \cdot \gamma^s) + \ln Y_i + \ln Y_j + (1-\sigma)\ln \tau_{ij} + (\sigma-1)\ln \beta_{ij}^s + (\sigma-1)\ln P_{f_j^s} + \ln(\Omega_i) \quad , \tag{2.17}$$

where $\ln X_{ij}^s$ is the logarithm of the exports from country i to country j in sector s, $\ln(\mu \cdot \gamma^s)$ is the logarithm of the income share spent in the (foreign) composite good s, $\ln Y_i$ and $\ln Y_j$ are the logarithms of the exporting country's income and the importing country's income, respectively, $\ln \tau_{ij}$ is the logarithm of the trade costs between country i and country j, $\ln \beta_{ij}^s$ is the logarithm of the to country j's preference bias toward the representative commodity produced by country i in sector s, $\ln P_{f_j^s}$ is the logarithm of the price index of the (foreign) composite good s, and $\ln \Omega_i$ represents the logarithms of the multilateral effects, a indicator that include all the characteristics of the country i's potential commercial partners (excluding country j).

We begin our empirical exercise with estimating the canonical gravity model. Then we progressively augment the baseline specification by introducing several fixed effects according to the dimension of the panel to properly control for the *multilateral* characteristics defined by Equation (2.17). The baseline specification is defined as follows:

$$\begin{aligned} \ln T_{ijst} \quad &= \quad \beta_0 + \beta_1 \ln GDP_{it} + \beta_2 \ln GDP_{jt} + \beta_3 \ln dist_{ij} + \beta_4 border_{ij} + \\ &\quad +\alpha_i + \eta_j + \delta_s + \ \theta_t + \varepsilon_{ijst} \\ &\quad i = 1, ..., 10,\ j = 1, ...180\ (i \neq j),\ s = 1, ..., 6,\ t = 1999, ...2011\ , \end{aligned} \tag{2.18}$$

where i is the reporter country, j is the partner country, s is the sector, and t is the year. The dependent variable ($\ln T_{ijst}$) is the logarithm of the annual exports (X_{ijst}) or imports (M_{ijst}) in sector s and year t (in current euro).[26] This dependent variable is expected to be affected

[26] The sample has a potential of $139,620$ observations (that is, $10 \cdot 179 \cdot 13 \cdot 6 = 139,620$). Concerning the

by a constant (β_0),[27] the logarithm of the reporter's GDP in year t ($\ln GDP_{it}$), the logarithm of the partner's GDP in year t ($\ln GDP_{jt}$), the logarithm of the distance between the reporter and the partner ($\ln dist_{ij}$), a dummy variable that indicates if the reporter and partner share a common border ($border_{ij}$), and the error term (ε_{ijst}). These explanatory variables are part of the standard gravity model and control for the countries' sizes and for the trade costs between any two commercial partners. The signs of the coefficients β_1, β_2, and β_4 are expected to be positive, indicating that the intensity of trade flows between two commercial partners increases with the size of their economies as well as their cultural links (measured by the *border* dummy variable), while β_3 is expected to be negative, indicating that the intensity of trade flows between two commercial partners decreases with their geographical distance. Moreover, in this econometric specification, we control for unobserved individual characteristics, namely fixed effects of the reporter (α_i), partner (η_j), sector (δ_t), and time (θ_t).[28] As we said before, these individual fixed effects are not capturing the *multilateral* effects of our theoretical model. But the objective of Equation (2.18) is to easily describe the two econometric strategies.

With the goal of determining the impact of the fifth EU enlargement on the intensity and direction of the EU-10's trade flows, we introduce in the baseline specification (2.18) a select group of DVs based on the DID strategy. This strategy is typically used to measure the causal effect of implementing a specific program or policy in a target group. In our case, the fifth EU enlargement represented a larger single market, completing a set of 25 countries that exchange goods and services without tariffs or quantitative controls. Our target group is those countries comprising the EU-10, and we are interested in measuring the effectiveness of the EU enlargement for this specific group. One canonical feature of the DID strategy is the choice of a temporal break, usually related to the time the policy is implemented. We thus define two complementary exercises. In the former, the temporal break corresponds to the year of the fifth EU enlargement (2004), and on the latter we do not make such a choice.

Our first empirical exercise proposes a pure DID strategy with the objective of determining whether or not the new EU member (the EU-10) experienced a change in intensity of their trade flows after becoming part of the existing single market. After defining our temporal periods and

---

trade flow data, there are a lot of missing values (37 percent in the case of exports and 43 percent in the case of imports), but the number of zeros is small (233 in the case of exports and 456 in the case of imports). In Sections 2.5 and 2.6 we address the zero values problem by adding one unit to the trade flow data before taking logarithms (Chen, 2004), while missing values are excluded from the sample (Harris et al., 2012). We need to take care of the potential bias that the exclusion of missing values might have on estimations (Gleditsch, 2002).

[27] For the sake of simplicity lnA in Equation (2.17) has been replaced by β_0.

[28] In this study, time fixed effects control for economic events that affect all countries and sectors, such as a financial crisis.

four partner groups we created several DVs: $af04$ equals one when the year is 2004 or later; $eu15$ equals one when the partner belongs to the EU-15; $eu10$ equals one when the partner belongs to the EU-10; and fsu equals one when the partner belongs to the FSU. We interact the temporal DV with the three partner DVs to generate commercial partner DVs after 2004, namely, $eu15_{af04}$, $eu10_{af04}$, and fsu_{af04}.[29] These DVs are added to the baseline specification (2.18) to get the following expression:

$$\begin{aligned}\ln T_{ijst} &= \beta_0 + \beta_1 \ln GDP_{it} + \beta_2 \ln GDP_{jt} + \beta_3 \ln dist_{ij} + \beta_4 border_{ij} + \\ &+ \gamma^{af04}_{eu15} eu15_{af04} + \gamma^{af04}_{eu10} eu10_{af04} + \gamma^{af04}_{fsu} fsu_{af04} + \\ &+ \gamma_{eu15} eu15 + \gamma_{eu10} eu10 + \gamma_{fsu} fsu + af04 + \alpha_i + \eta_j + \delta_s + \theta_t + \varepsilon_{ijst} \qquad (2.19)\end{aligned}$$

$$i = 1, ..., 10,\ j = 1, ...180\ (i \neq j),\ s = 1, ..., 6,\ t = 1999, ...2011\ .$$

The countries belonging to the ROW and the period 1999-2003 are the references (or control groups) used to interpret the econometric results. Therefore, the coefficient γ^{af04}_{g} (g = *eu15*, *eu10*, or *fsu*) in Equation (2.19) can be expressed as follows:

$$\gamma^{af04}_{g} = \left(\overline{T}_{g,\ af04} - \overline{T}_{g,\ bef04}\right) - \left(\overline{T}_{\text{row}\ ,\ af04} - \overline{T}_{\text{row}\ ,\ bef04}\right)\ , \qquad (2.20)$$

where $\overline{T}$ represents the annual average level of trade flows (exports or imports) for the corresponding period ($bef04$ refers to the period 1999–2003, while $af04$ refers to the period 2004–2011). In other words, the coefficient γ^{af04}_{g} describes the increase in the exports (or imports) between the two defined periods by comparing group g versus the ROW. A positive sign indicates that the intensity of the EU-10's export flows (or import flows) increased to (from) group g more than it did to (from) the ROW, which could be interpreted as *trade redirection* from the ROW to group g.[30] *A priori* γ^{af04}_{eu15} and γ^{af04}_{eu10} are expected to be positive and γ^{af04}_{fsu} is expected to be negative. In other words, we expect trade redirection from the ROW to the EU countries (EU-15 and EU-10) because they are part of the same economic community, and we expect no such effect to the FSU.

[29] For instance, the DV $eu15_{af04}$ equals one when the partner belongs to the EU-15 and the year is 2004 or later.

[30] It is important to take into account that this coefficient must be interpreted as differences and not as levels. If $\gamma^{af04}_{g} > 0$, we cannot claim that the level of exports to group g was higher than the level of exports to the ROW, yet we can claim that the increase in export flows to group g was higher than the increase in export flows to the ROW.

Because trade flow data is reported annually, 2004 (the year of the fifth EU enlargement) has to be included in either the treatment or control group. In the first exercise, it is included in the treatment group.[31] To test whether this decision could determine the results, we propose a second exercise where each commercial partner DV (*eu15*, *eu10*, and *fsu*) interacts with the temporal DVs, now defined for each year. Using 1999 as the time reference, we can study the tendency of each group throughout the entire period, year by year.[32] To do this, we introduce in the baseline specification (2.18) a new set of DVs, and the equation becomes:

$$\ln T_{ijst} = \beta_0 + \beta_1 \ln GDP_{it} + \beta_2 \ln GDP_{jt} + \sum_{t=2000}^{2011} \gamma^t_{eu15} eu15^t + \sum_{t=2000}^{2011} \gamma^t_{eu10} eu10^t + \sum_{t=2000}^{2011} \gamma^t_{fsu} fsu^t + \lambda_{ijs} + \theta_t + \varepsilon_{ijst}$$

$$i = 1, ..., 10,\ j = 1, ...180\ (i \neq j),\ s = 1, ..., 6,\ t = 1999, ...2011\ , \qquad (2.21)$$

where $eu15^t$, $eu10^t$, and fsu^t are the commercial partner DVs in year t and λ_{ijst} is the three-dimensioned fixed effects (at the reporter–partner–sector level).[33]

This technique allows us to compare the changes in the intensity of trade flows between the EU-10 and each commercial partner group g against the changes in the intensity of trade flows between the EU-10 and the ROW for any period of time. For example, the coefficient of the partner group g multiplied by the year dummy $year03$ is equal to:

$$\gamma_g^{2003} = \left(\overline{T}_{g,\,2003} - \overline{T}_{g,\,1999}\right) - \left(\overline{T}_{\text{row}\,,\,2003} - \overline{T}_{\text{row}\,,\,1999}\right) \quad , \qquad (2.22)$$

where $\overline{T}$ represents the annual average level of trade flows (exports or imports). The coefficient γ_g^{2003} describes the change in the intensity of trade flows between 1999 and 2003 between two groups, group g and the ROW. A positive (negative) sign indicates that the trade variation was higher (lower) for group g than for the ROW.

This technique is more flexible than the usual DID strategy in the way it manages the temporal dimension. In addition, it emphasizes the precise moment at which the trade variation took place, using 1999 as the reference. Henceforth, we refer to the first econometric specification as the DID strategy defined by Equation (2.19), and the second as the trend of the commercial partner groups over time defined by Equation (2.21).

[31] Hornok (2010) tackles the mid-year accession problem by considering only the odd years. In this case, the treatment group includes the years 2005 and 2007, and the control group includes the years 1999, 2001, and 2003.

[32] Carrère (2006) and Head et al. (2010) propose a similar exercise to study the effects of regional trade agreements and former colonies, respectively, on the intensity of trade flows along a period of time.

[33] Because we consider reporter–partner–sector fixed effects, we indirectly consider the effect of each commercial partner group for the entire period. Therefore, we were forced to use $eu15^{1999}$, $eu10^{1999}$, and fsu^{1999} as control groups.

2.6 Results

To present the results clearly, this section is divided into two parts. The first focuses on the standard DID strategy and the second on the trend analysis. For the first time, we study the effects of the EU accession on both the EU-10's export and import flows. Changes in export flows usually rely on competitiveness factors (Antimiani and Costantini, 2013), while changes in import flows are associated with trade creation and trade diversion mechanisms created by the existence of a regional trade agreement (Carrère, 2006). We compute, in each exercise, the effects based on all the sectors jointly and again with each sector individually. The latter exercise allows us to identify the existence of sector-specific effects that could be hidden if the analysis were done only on all sectors jointly.

2.6.1 Difference-in-Differences Strategy

The objective of this estimation exercise is to assess whether or not there was variation in the intensity of the EU-10's trade flows after 2004. To this end, we split the time period into two parts and compare the level of the EU-10's trade flows between the two periods for each of the four commercial partner groups, as described by Equation (2.19).

Focusing on export flows and taking all sectors jointly, the results in Table 2.1 are presented according to the way heterogeneity has been controlled, by including various fixed effects and interactions. In column 1, the estimation is run by OLS and all individual fixed effects are considered—that is, reporter (α_i), partner (η_i), sector (δ_i), and year (θ_t). Column 2 includes the reporter–partner fixed effects (μ_{ij}). Because these fixed effects capture all the time invariant characteristics, some variables ($\ln dist_{ijt}$ and $border_{ijt}$) are dropped.[34] Finally, column 3 includes the reporter-partner-sector fixed effects (λ_{ijs}). The model is run using the fixed-effects (FE) or within transformation estimator. In all cases, the standard errors (SEs) are clustered at the reporter-partner level, then some dependence among country-pairs is allowed.

We discuss only the results in column 3, which is our preferred specification. The coefficients of the GDP are positive and statistically significant. If the exporting country's GDP increases by one percent, the export flows increase by 0.87 percent; if the importing country's GDP increases by one percent, the export flows increase by 0.68 percent. As expected, the coefficients $\widehat{\gamma}_{eu15}^{af04}$ and $\widehat{\gamma}_{eu15}^{af04}$ are positive and statistically significant, meaning that after 2004 the variation of the

[34] To run this regression we employed the program *reg2hdfe* in Stata, implementing the algorithm developed by Guimarães and Portugal (2010). This program allows us to control for high-dimensional fixed effects (in our case reporter–partner) without incurring in storage problems. But it does not report a constant.

intensity of the EU-10's export flows to the EU-15 and EU-10 increased more than that to the ROW. Specifically, the EU-10's export flows to the EU-15 and EU-10 increase by 14.85 percent and 23.93 percent, respectively, versus the reference group.[35] If we compare these two coefficients we cannot reject the null hypothesis that their difference is equal to zero. Therefore, we can conclude that after 2004 there was a trade integration process toward the EU, meaning that the EU-10 found EU markets attractive for selling their products. Furthermore, the coefficient $\widehat{\gamma}_{fsu}^{af04}$ is not statistically different from zero, meaning that after 2004, the variation of the EU-10's export flows to the FSU was not different from that to the ROW.

From an econometric viewpoint two comments must be made. First, the introduction of fixed effects that account for more than one dimension affects the magnitude and significance of the coefficients of interest, namely $\widehat{\gamma}_{eu15}^{af4}$, $\widehat{\gamma}_{eu10}^{af04}$, and $\widehat{\gamma}_{fsu}^{af04}$. The last of these is positive and statistically significant in the first two specifications but not in the last one. Second, column 3 reports a constant, but its interpretation is not very intuitive. The within transformation model considers more than $9,000$ unobserved individual effects, one for each reporter–partner–sector combination (there are $9,369$ groups), and an intercept is estimated for each reporter–partner–sector combination ($\widehat{\lambda}_{ijs}$). The constant reported by the FE model ($\widehat{\beta}_0$) is the average of them all (Wooldridge, 2006), but also captures the average values of the quantitative regressors ($lnGDP_{it}$ and $lnGDP_{jt}$) and the effects of the control groups of the DVs (namely, row, $year99$, and $year04$).

Another interesting exercise is to assess whether or not the effects of the fifth EU enlargement on the intensity and direction of the EU-10's export flows were homogeneous across all sectors. Table 2.2 reports the results for the FE model by sector. In the chemicals and F&B sectors, we see trade redirection from the ROW to the other partner groups. The greatest impact in the F&B exports was to the EU-15.[36] The machinery and vehicles sector (*Machinery*) is one of the most heterogeneous sectors in the sense that it is composed of several types of products of different qualities. The estimation referring to this sector emphasizes that after 2004, the EU-10 experienced export redirection from the EU-15 to the ROW, meaning that the intensity of the EU-10's export flows to the ROW increased more than that to the EU-15. An explanation could be related to competitiveness: the quality of the products from the EU-10 might not fit

[35] For example, $\widehat{\gamma}_{eu15}^{af04} = 0.1385$ indicates that the increase in exports is equal to $[\exp(0.1385) - 1] \cdot 100 = 14.85$ percent (Halvorsen and Palmquist, 1980).

[36] The coefficient $\widehat{\gamma}_{eu15}^{af04}$ is statistically different from the other two coefficients ($\widehat{\gamma}_{eu10}^{af04}$ and $\widehat{\gamma}_{fsu}^{af04}$) at the one percent level. According to the EU official webpage (http://europa.eu/index_en.htm), the Czech Republic produces a world-famous beer (namely, Pilsner) and wine. The Hungarian wines (for example, Tokaji) are also known worldwide.

Table 2.1: EU-10 export flows

Dependent variable:	$\ln(\text{exports}_{ijst}+1)$		
	(1)	(2)	(3)
Estimator:	OLS	OLS	FE
lnGDP_{it}	0.6859***	0.7515***	0.8719***
	(0.1111)	(0.1089)	(0.1083)
lnGDP_{jt}	0.5381***	0.6033***	0.6764***
	(0.0619)	(0.0610)	(0.0608)
lndist_{ij}	-2.1620***		
	(0.1077)		
border_{ij}	0.6205***		
	(0.1654)		
eu15_{af04}	0.2119***	0.1964***	0.1385***
	(0.0502)	(0.0501)	(0.0502)
eu10_{af04}	0.2935***	0.2735***	0.2146***
	(0.0691)	(0.0684)	(0.0679)
fsu_{af04}	0.1910**	0.1700**	0.1188
	(0.0789)	(0.0797)	(0.0823)
constant	1.5831		-25.1250***
	(3.2604)		(2.9830)
Fixed effects:			
reporter (r)	x		
partner (p)	x		
sector (s)	x	x	
time	x	x	x
r-p		x	
r-p-s			x
Observations	87195	87195	87195
Number of groups			9369
R-sq	0.6273	0.6870	
R-sq (overall)			0.3506
R-sq (between)			0.4300
R-sq (within)			0.1222
F	131.0451		196.5873
pvalue	0.0000		0.0000
sigma_u			2.8049
sigma_e			1.5503
rho			0.7660

Standard errors clustered by reporter–partner are in parenthesis.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Specification (1) includes the DVs $eu15$, $eu10$, fsu, and $af04$.
Specifications (2) and (3) include the DV $af04$.

Table 2.2: EU-10 export flows by sector

	Dependent variable: ln(exports$_{ijt}$+1) - Estimator: FE					
	Chemicals	F&B	Machinery	Manufbymat	Othermanuf	Rawmat
lnGDP$_{it}$	0.3312	0.1487	1.9443***	0.4903**	1.2729***	0.3354
	(0.2276)	(0.2635)	(0.1756)	(0.2113)	(0.2169)	(0.2806)
lnGDP$_{jt}$	0.6046***	0.2358*	0.8182***	0.6956***	0.7543***	0.9544***
	(0.1128)	(0.1423)	(0.1043)	(0.1154)	(0.1220)	(0.1453)
eu15$_{af04}$	0.5447***	1.1560***	-0.2480***	0.1620**	-0.3538***	-0.4051***
	(0.1162)	(0.1240)	(0.0795)	(0.0823)	(0.0879)	(0.1232)
eu10$_{af04}$	0.3214*	0.7485***	-0.1585	0.5386***	-0.0713	-0.0387
	(0.1679)	(0.1532)	(0.1109)	(0.1210)	(0.1125)	(0.1564)
fsu$_{af04}$	0.2588**	0.3999**	0.2213	0.4977***	-0.2457*	-0.3849**
	(0.1220)	(0.1576)	(0.1450)	(0.1571)	(0.1451)	(0.1946)
constant	-10.6567*	2.8638	-53.4858***	-15.9440***	-37.0651***	-19.9111***
	(6.1788)	(7.0843)	(4.9601)	(5.9512)	(5.7887)	(7.3706)
Observations	14783	12725	16624	15357	16074	11632
Number of groups	1579	1502	1684	1605	1639	1360
R-sq (overall)	0.3509	0.2373	0.4252	0.4226	0.3662	0.2313
R-sq (between)	0.4264	0.3081	0.5308	0.5324	0.4726	0.2628
R-sq (within)	0.1629	0.0773	0.2336	0.0508	0.1286	0.1726
F	64.2422	31.4486	113.6239	25.8378	55.2661	52.8541
pvalue	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
sigma_u	2.7705	3.0871	2.5345	2.8864	2.6771	2.9600
sigma_e	1.4272	1.6625	1.4392	1.5060	1.5697	1.6434
rho	0.7903	0.7752	0.7562	0.7860	0.7442	0.7644

Standard errors clustered by reporter–partner are in parenthesis.

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

Fixed-effects dimension: reporter–partner (μ_{ij}). All regressions include time DVs ($af04$ and θ_t).

well with the preferences of the EU-15, causing the EU-10 to search for new markets. This argument is supported by the evidence discussed in Head and Mayer (2004), where trade is found to be more attractive between countries with similar GDPs because citizens' preferences are more comparable. Exports of manufactured goods classified by material (*Manufbymat*) were redirected from the ROW to the other groups. Exports of other manufactured articles (*Othermanuf*) were redirected from the EU-15 and FSU to the ROW. The argument put forth for the machinery and vehicle sector holds for this sector as well, which includes goods of high technological worth.[37] Finally, exports of raw materials and energy (*Rawmat*) were redirected from the EU-15 and FSU to the ROW. This is not surprising because Russia, the largest country of the FSU, is a major exporter of gas, and the EU is a major importer of Russian natural gas (Noël, 2008).

Hence, our results suggest that the effects of the fifth EU enlargement on the EU-10's export flows to the EU-15 were not homogeneous across all sectors. We find that after 2004, the EU-10's export flows to the EU-15 increased more than those to the ROW in some sectors (chemicals, F&B, and manufactured goods by material), but increased less in others (machinery and vehicles, other manufactured articles, and raw materials and energy).

To complete our first estimation exercise, we apply the same empirical analysis to the EU-10's imports. First, all sectors are analyzed jointly (Table 2.3). Again, we focus on the results in the third column, containing the FE model. The coefficients of the GDP are positive and statistically significant. If the importing country's GDP increases by one percent, the import flows increase by 0.62 percent; if the exporting country's GDP increases by one percent, the import flows increase by 0.24 percent. We observe that the increase in the intensities of the EU-10's import flows from the EU-15, EU-10, and FSU are higher than those from the ROW. Nevertheless, the most important change occurred from the EU-10 ($\widehat{\gamma}_{eu10}^{af04} = 1.22$).[38] Therefore, trade liberalization had an important impact on internal EU-10 trade flows, which could be related to time-variant bias in EU-10 consumer preferences; the EU-10's buyers might have perceived EU-10 products *ipso facto* to be different from non-EU-10 products. Hence, buyers preferred to import commodities from EU-10 countries instead of from EU-15 countries, which illustrates a group-biased demand effect.[39] Unlike that seen among exports, the incorporation

[37] For example, Division 87 of the SITC comprises professional, scientific and controlling instruments and apparatus, n.e.s., while Division 88 comprises photographic apparatus, equipment and supplies and optical goods, n.e.s.; watches and clocks (UN, 2006).

[38] The coefficient $\widehat{\gamma}_{eu10}^{af04}$ is statistically different from the other two coefficients ($\widehat{\gamma}_{eu15}^{af04}$ and $\widehat{\gamma}_{fsu}^{af04}$) at the one percent level.

[39] Head and Mayer (2000) use the home-biased demand assumption to explain EU market fragmentation in the mid-1980s.

of fixed effects that captures more than one dimension did not affect the magnitude of the coefficients of interest as much.

To assess whether or not the variation of the intensity of the EU-10's import flows was common among all sectors, we propose sectoral analysis (Table 2.4). It is important to remark that the capacity of prediction of the model for changes in the EU-10's import flows (by sector) is more limited than the one for exports. We find that the EU-10 increased its import flows more from the FSU than from the ROW in chemicals, F&B, and other manufactured articles sector, while the EU-10 increased its import flows more from the EU-15 and EU-10 than from the ROW in all sectors. In addition, except for the F&B, the highest variation was within the EU-10, meaning that group-biased demand was a generalized effect that extended to almost all sectors.

Table 2.3: EU-10 import flows

Dependent variable:	$\ln(\text{imports}_{ijst}+1)$		
	(1)	(2)	(3)
Estimator:	OLS	OLS	FE
lnGDP_{it}	0.3732***	0.3870***	0.6169***
	(0.1263)	(0.1255)	(0.1213)
lnGDP_{jt}	0.2260***	0.2326***	0.2440***
	(0.0702)	(0.0708)	(0.0681)
lndist_{ij}	-1.9184***		
	(0.1040)		
border_{ij}	0.6932***		
	(0.1889)		
eu15_{af04}	0.5943***	0.6078***	0.6963***
	(0.0480)	(0.0484)	(0.0491)
eu10_{af04}	1.0733***	1.0998***	1.2158***
	(0.0878)	(0.0884)	(0.0864)
fsu_{af04}	0.4839***	0.4672***	0.4509***
	(0.0986)	(0.1022)	(0.0951)
constant	9.7911***		-8.5088**
	(3.5842)		(3.2999)
Fixed effects:			
reporter (r)	x		
partner (p)	x		
sector (s)	x	x	
time	x	x	x
r-p		x	
r-p-s			x
Observations	79848	79848	79848
Number of groups			9003
R-sq	0.6715	0.7099	
R-sq (overall)			0.2964
R-sq (between)			0.3961
R-sq (within)			0.0355
F	285.2938		75.6782
pvalue	0.0000		0.0000
sigma_u			3.9221
sigma_e			1.5972
rho			0.8577

Standard errors clustered by reporter–partner are in parenthesis.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Specification (1) includes the DVs $eu15$, $eu10$, fsu, and $af04$.
Specifications (2) and (3) include the DV $af04$.

Table 2.4: EU-10 import flows by sector

	Dependent variable: ln(imports$_{ijt}$+1) - Estimator: FE					
	Chemicals	F&B	Machinery	Manufbymat	Othermanuf	Rawmat
lnGDP$_{it}$	0.0660	0.1848	1.0004***	1.1158***	0.9785***	0.1472
	(0.2630)	(0.2311)	(0.2407)	(0.2116)	(0.2026)	(0.2695)
lnGDP$_{jt}$	-0.0969	0.2738**	0.6156***	0.1773	0.3341***	0.0135
	(0.1568)	(0.1341)	(0.1330)	(0.1384)	(0.1175)	(0.1518)
eu15$_{af04}$	0.3488***	1.2846***	0.4000***	0.4585***	0.8014***	0.6985***
	(0.0927)	(0.0903)	(0.0834)	(0.0740)	(0.0692)	(0.1200)
eu10$_{af04}$	1.3801***	1.5234***	0.9277***	1.0230***	1.0568***	1.2715***
	(0.1689)	(0.1550)	(0.1433)	(0.1456)	(0.1272)	(0.1771)
fsu$_{af04}$	0.7397***	0.8309***	0.0523	0.3515	0.7604***	0.0038
	(0.2309)	(0.1905)	(0.2003)	(0.2181)	(0.1909)	(0.2277)
constant	13.1428*	1.5326	-26.6676***	-18.7123***	-19.8917***	8.3681
	(7.3710)	(6.4755)	(6.6953)	(6.0676)	(5.5840)	(7.3220)
Observations	10726	14081	13821	13520	15444	12256
Number of groups	1294	1537	1586	1522	1637	1427
R-sq (overall)	0.0008	0.3936	0.3896	0.0905	0.2583	0.1112
R-sq (between)	0.0118	0.5209	0.4890	0.1484	0.3630	0.1976
R-sq (within)	0.0696	0.0522	0.0507	0.0314	0.0341	0.0368
F	27.5769	30.6800	33.2839	21.3612	23.3854	19.4571
pvalue	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
sigma_u	4.6799	3.3640	3.7842	4.2857	3.8846	3.7372
sigma_e	1.5430	1.4915	1.7180	1.5796	1.4794	1.7341
rho	0.9020	0.8357	0.8291	0.8804	0.8733	0.8228

Standard errors clustered by reporter–partner are in parenthesis.

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

Fixed-effects dimension: reporter–partner (μ_{ij}). All regressions include time DVs ($af04$ and θ_t).

2.6.2 The Trend of the Commercial Partner Groups over Time

This additional empirical exercise aims at overcoming the time issue that could not be addressed by the usual DID strategy. The DID strategy relies heavily on the definition of a temporal control group, which by construction, is defined as 1999–2003. Nevertheless, if entrance to the single market had an anticipated trade effect, the period 1999–2003 might not be an adequate control. In this exercise, we fix the initial year of the period (1999) as the reference point, and observe the variation in the intensity of trade flows between the EU-10 and its commercial partners over time, using the ROW as the partner reference, as shown in Equation (2.21). In this way, we can confirm whether or not the results obtained from the DID strategy are consistent or sensitive to the definition of the temporal control group.

The results are presented graphically to summarize and favor their interpretation. The results of the regressions that consider all sectors jointly are illustrated in Figures 2.2 and 2.3, for exports and imports, respectively. These regressions are estimated using reporter–partner–sector fixed effects (λ_{ijs}). The graphs depict the values of the coefficients of interest over the period 2000–2011 (γ^t_{eu15}, γ^t_{eu10}, and γ^t_{fsu}, $t = 2000, ..., 2011$), as well as their confidence intervals at the five percent level.[40] The sectoral analysis is included in Appendix 2.9 (Figures 2.6–2.11). In this case, the fixed effects are at reporter–partner level (μ_{ij}).

We analyze the EU-10's export flows to each commercial partner group separately (Figure 2.2). We find that from 2004 to 2010 there are differences between exports flows to the EU-15 and those to the ROW (control group), evidenced by the positive and statistically significant coefficient γ^t_{eu15}. The coefficient does not follow any specific trend; thus we can assume that the redirection of the EU-10's export flows was an isolated event that occurred in the year of the fifth EU enlargement (2004).[41] Figure 2.6 in Appendix 2.9 shows the same conclusions for chemicals and F&B sectors, but a negative or null effect for all other sectors. These results support the heterogeneous behavior in the variation of the intensity of the EU-10's export flows to the EU-15 previously detected. We find exports to the EU-10 increased more than those to the ROW from 2006 to 2011, evidenced by the positive and statistically significant coefficient $\widehat{\gamma}^t_{eu10}$. In contrast, the coefficient $\widehat{\gamma}^t_{fsu}$ is negative and statistically significant from 2001 to 2004, suggesting that there was export redirection from the FSU to the ROW prior to the fifth EU enlargement. At the sectoral level this pattern is also found in the chemicals sector (see Figure

[40] A coefficient is statistically significant if the range of its confidence interval does not contain the value zero. Graphically this means that the horizontal line $y = 0$ does not cross the confidence interval, which is represented in grey.

[41] The variation between two consecutive coefficients is not statistically different from zero except between 2008 and 2009, when negative changes were caused by economic crisis.

2.8 in Appendix 2.9). This outcome encourages us to refine the conclusions reported in the previous subsection, where we find that the fifth EU enlargement did not impact the intensity of the EU-10's exports to the FSU.

Finally, we examine the EU-10's import flows (Figure 2.3). We find that from 2004 to 2011, imports from the EU-15 increased more than those from the ROW, evidenced by the positive and statistically significant coefficient $\widehat{\gamma}^t_{eu15}$. In this case, trade redirection was not isolated; we identify a positive trend from 2003 to 2006 and from 2008 to 2010.[42] Meanwhile, the coefficient $\widehat{\gamma}^t_{eu10}$ is positive and statistically significant from 2004 to 2011, following a slight positive trend (except from 2006 to 2007, where it is constant). In Figure 2.10 (Appendix 2.9), we present evidence for the existence of a positive change from 2003 to 2004 for all sectors. The coefficient $\widehat{\gamma}^t_{fsu}$ is positive and statistically significant from 2005 to 2011 (except in 2009), a pattern that is also found in the F&B sector (see Figure 2.11 in Appendix 2.9).

We can highlight several interesting findings resulting from this second empirical exercise. For one, we can confirm that the redirection of the EU-10's exports from the ROW to the EU-15 was an isolated event occurring in a few sectors in 2004. We also find a negative change of the intensity of exports to the FSU immediately prior to the official entry of the EU-10 into the EU. Additionally, we find a huge variation of the intensity of the EU-10's import flows from the EU-10 from 2003 to 2004 in all sectors.[43] Unlike the DID strategy, this technique shows the trend of each commercial partner group during the entire period, using 1999 as the reference. In general, the effects on the EU-10's imports were stronger than those on their exports.

[42] This means that the variation between two consecutive coefficients is positive and statistically significant.

[43] The coefficient $\widehat{\gamma}^t_{eu10}$ is statistically different from $\widehat{\gamma}^t_{eu15}$ in all sectors except F&B. In Appendix 2.9, Figure 2.9 shows the trend of the latter coefficient over time and Figure 2.10 shows the same for the former coefficient.

Figure 2.2: Trend of the commercial partner DVs over time: EU-10 exports

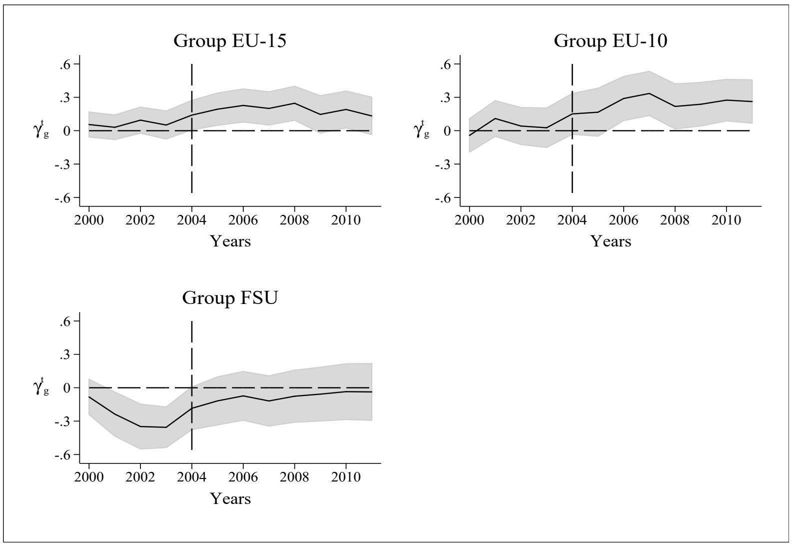
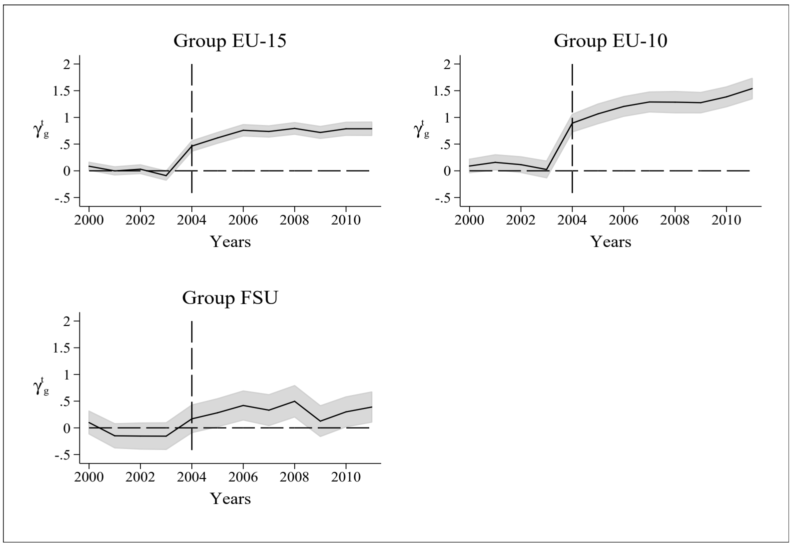


Figure 2.3: Trend of the commercial partner DVs over time: EU-10 imports



2.7 Extensions

In this section, we implement two additional exercises to check the consistency of the results discussed in Section 2.6. In the former exercise, we deal with the missing values problem in a different way. Now, missing values are not excluded from the sample and they are replaced by zeros. Then, an alternative estimator method is implemented in order to deal with the large number of zero values. In the latter exercise, an additional explanatory variable is included in the baseline econometric specification (2.18).

Although one of the most important concerns of the trade flow literature is how to deal with missing values and zeros, there is not a consensus about how to handle this important issue. In some studies, missing values are treated as zeros (Brun et al., 2005; Coe et al., 2007; Felbermayr and Kohler, 2006; Gleditsch, 2002), while others insists that these values are conceptually different (Harris et al., 2012). In Sections 2.5 and 2.6 we address the zero values problem by adding one unit to the trade flow data before taking logarithms (Chen, 2004), while missing values were excluded from the sample. Nevertheless, excluding the missing observations is not riskless. Gleditsch (2002) discusses and demonstrates that the removal of missing values could lead to non-random samples that produces misleading inferences and results. King et al. (2001) confirm that the consequences from the missing values can be worse than those from the omitted variable bias problem. Gleditsch (2002) proposes several procedures to solve the problem, one among others to replace the missing values by zeros. He argues that when a trade flow data reported by two commercial partners are missing, it means that the intensity of trade flows among these countries is expected to be small or negligible. Following this reasoning, as a first robustness check, we assume the most extreme working hypothesis and we consider that a missing value is equal to a zero value. In this way, it is possible to test whether or not excluding missing values created a biased sample.

When equalizing missing values to zero values, we need to face the necessity of finding a robust estimator for a sample where the dependent variable has a high proportion of zeros. The literature suggests exploiting the Poisson distribution (Santos Silva and Tenreyro, 2006), which is usually implemented in count data models, where the dependent variable can take nonnegative values (zeros and positive values). In our case, the value of the exports (or imports) is the response variable. The Poisson regression shapes the logarithm of the expected count as a function of a group of selected explanatory variables. For instance, the coefficient γ_g^{af04} of Equation (2.19) indicates the difference (in units) in the logarithms of the expected counts between group g and the ROW after 2004 (assuming that the other variables are constant). Following Westerlund and Wilhelmsson (2011) we implemented a FE Poisson ML estimator. This model allows us to correct not only for the zero values (or missing values) but also for the heteroskedaticity inherent in log-linearized models (Santos Silva and Tenreyro, 2006).

Because the coefficients reported by the FE Poisson ML model are not directly comparable with those obtained from the FE model (Tables 2.1–2.4), we focus primarily on the significance of the coefficients of interest without making cross comparisons. Table 2.5 shows the results for all sectors jointly; column 1 for the exports and column 2 for the imports. Among exports, the coefficients $\widehat{\gamma}_{eu15}^{af04}$ and $\widehat{\gamma}_{eu10}^{af04}$ are not statistically different from zero while they are in the FE model (Table 2.1, column 3). Among imports, these coefficients are statistically significant at the one percent level as they are in the FE model (Table 2.3, column 3). This new estimator confirms that there was a redirection of the EU-10's imports from the ROW to the EU markets after 2004, above all toward the EU-10 market.

We use the sectoral analysis to check the validity of our results at a more disaggregated level (see Table 2.6 for exports by sector and Table 2.7 for the imports by sector). Among exports, some of the results are consistent with those found in the FE model (Table 2.2). For instance, the coefficient $\widehat{\gamma}_{eu15}^{af04}$ is positive in two of the sectors and negative in three. Among imports (Table 2.7), the coefficients $\widehat{\gamma}_{eu15}^{af04}$ and $\widehat{\gamma}_{eu10}^{af04}$ are positive and statistically significant for four and five sectors, respectively.[44] Then, we can declare that our main results are robust to the implementation of another estimator method that handles the problem of the missing and zero values.

[44]Table 2.7 does not report results for the machinery and vehicles sector because the FE Poisson ML estimator did not converge. We attempt to fix this technical problem by following the suggestions of Santos Silva and Tenreyro (2011), but unfortunately they do not apply to our case because we have a panel of data. One of the possible sources of the problem is the existence of perfect collinearity between regressors in the sample where the dependent variable (the imports) is positive.

Table 2.5: EU-10 trade flows: extension 1

Estimator:	FE Poisson ML	
	(1)	(2)
Dependent variable:	exports_{ijst}	imports_{ijst}
lnGDP_{it}	0.5649***	0.6773***
	(0.1138)	(0.1072)
lnGDP_{jt}	0.7418***	0.5635***
	(0.0730)	(0.0731)
eu15_{af04}	-0.0392	0.2694***
	(0.0538)	(0.0617)
eu10_{af04}	0.0328	0.4541***
	(0.0613)	(0.0736)
fsu_{af04}	-0.0269	-0.1132
	(0.0833)	(0.0953)
Observations	120694	115362
Wald	4384.4776	5251.6186
pvalue	0.0000	0.0000
iterations	5	4

SEs clustered by reporter–partner–sector are in parenthesis.

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

Fixed-effects dimension: reporter–partner–sector (λ_{ijs}).

All regressions include time DVs ($af04$ and θ_t).

Table 2.6: EU-10 export flows by sector: extension 1

	Dependent variable: exports_{ijt} - Estimator: FE Poisson ML					
	Chemicals	F&B	Machinery	Manufbymat	Othermanuf	Rawmat
lnGDP_{it}	0.1806	0.5265**	0.9590***	0.0861	0.7211***	-0.2336
	(0.1814)	(0.2455)	(0.1945)	(0.0943)	(0.1701)	(0.2098)
lnGDP_{jt}	0.2161*	0.2968*	0.9587***	0.5908***	0.8559***	0.7641***
	(0.1116)	(0.1635)	(0.1351)	(0.1016)	(0.1084)	(0.1611)
eu15af04	0.2161**	0.7565***	-0.1015	0.0680	-0.1603**	-0.4700**
	(0.0897)	(0.1179)	(0.0927)	(0.0609)	(0.0765)	(0.2332)
eu10_{af04}	0.0698	0.6165***	0.1196	0.1388*	0.0777	-0.3966*
	(0.0899)	(0.1268)	(0.1270)	(0.0757)	(0.0975)	(0.2293)
fsu_{af04}	0.3874***	0.2067	0.0826	0.1598*	-0.2148	-0.7201***
	(0.1104)	(0.1609)	(0.1506)	(0.0928)	(0.1466)	(0.2602)
Observations	20334	19341	21703	20667	21132	17517
Number of groups	1576	1498	1682	1602	1638	1357
Wald	1700.1019	872.8663	2835.5171	7087.0012	1650.1746	1448.2662
pvalue	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
iterations	5	4	5	4	4	4

Standard errors clustered by reporter–partner are in parenthesis.

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

Fixed-effects dimension: reporter–partner (μ_{ij}). All regressions include time DVs ($af04$ and θ_t).

Table 2.7: EU-10 import flows by sector: extension 1

	Dependent variable: $imports_{ijt}$ - Estimator: FE Poisson ML				
	Chemicals	F&B	Manufbymat	Othermanuf	Rawmat
$\ln GDP_{it}$	0.3471***	0.3294**	0.6488***	1.0346***	0.0369
	(0.0992)	(0.1535)	(0.1102)	(0.1923)	(0.2346)
$\ln GDP_{jt}$	0.5173***	-0.2752*	0.4143***	0.3533***	0.0201
	(0.1159)	(0.1471)	(0.0933)	(0.1273)	(0.1019)
$eu15_{af04}$	0.4523***	0.9735***	0.0738	0.3020**	0.4522**
	(0.0965)	(0.1352)	(0.0881)	(0.1457)	(0.1775)
$eu10_{af04}$	0.5355***	1.1665***	0.2000**	0.5385***	0.5382***
	(0.0986)	(0.1356)	(0.0984)	(0.1056)	(0.1883)
fsu_{af04}	-0.0582	0.5862***	-0.2862*	-0.1455	0.2843
	(0.1140)	(0.2049)	(0.1678)	(0.1411)	(0.1779)
Observations	16538	19761	19486	20929	18381
Number of groups	1281	1530	1510	1622	1421
Wald	3357.1812	1443.2650	3558.8817	2696.3660	2624.6489
pvalue	0.0000	0.0000	0.0000	0.0000	0.0000
iterations	4	4	4.	4	5

Standard errors clustered by reporter–partner are in parenthesis.

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

Fixed-effects dimension: reporter–partner (μ_{ij}).

All regressions include time DVs ($af04$ and θ_t).

The estimator does not converge for machinery and vehicle sector.

Finally, we propose another extension to study the effects on the coefficients of interest when we include other relevant explanatory variables, such as FDI, in the baseline specification (2.18). According to the economic literature, FDI is considered complementary or substitutes for exports (Markusen, 2002; Pantulu and Poon, 2003). Under this perspective, there is the possibility that the intensity and direction of the EU-10's trade flows after 2004 could be associated with the potential entry or exit of FDI in the EU-10 countries. If FDI and trade flows were complementary, more FDI from country i to country j would positively affect the exports from country i to country j. On the contrary, if FDI and trade flows were substitutes, more FDI from country i to country j would have a negative effect on the exports from country i to country j.

We use FDI stock data, extracted from Eurostat database *EU Direct Investments - Main Indicators*. These data capture the size of the foreign stock of capital in the host economy.[45] In monetary terms, the EU-10's inward FDI stocks were much more considerable than the EU-10's outward FDI stocks. In 1999, the EU-10's inward FDI stock represented $300,510$ million euros and their outward FDI stock represented $144,150$ million euros. This difference sharply increased in 2011; inward FDI stock represented $2,446,158$ million euros and outward FDI stock represented $159,462$ million euros. Table 2.8 tabulates the EU-10's inward FDI stock. The main investor in the EU-10 market was the EU-15, representing around 80 percent of the total FDI. According to Lipsey (2006), the CEECs became one of the major locations for FDI from Europe, particularly from Germany, starting in 1990.

Table 2.8: EU-10 inward FDI stock, 1999 and 2011

	Year 1999		Year 2011	
investor	Million euros	%	Million euros	%
EU-15	243,288	80.96	1,961,172	80.17
EU-10	6,696	2.23	171,288	7.00
FSU	1,350	0.45	19,806	0.81
ROW	49,176	16.36	293,892	12.01
Total	300,510		2,446,158	

Source: Own elaboration based on data from Eurostat.

[45]The FDI stocks are measured according the value of the foreign investment at the end of the period.

The econometric specification for this extension becomes an augmented version of baseline econometric specification (2.18):

$$\begin{aligned}\ln T_{ijst} &= \beta_0 + \beta_1 \ln GDP_{it} + \beta_2 \ln GDP_{jt} + \beta_3 \ln FDI^{d}_{ijt} + \\ &+ \gamma^{af04}_{eu15} eu15_{af04} + \gamma^{af04}_{eu10} eu10_{af04} + \gamma^{af04}_{fsu} fsu_{af04} + af04 + \lambda_{ijs} + \theta_t + \varepsilon_{ijst} \qquad (2.23) \\ & i = 1, ..., 10,\ j = 1, ...180\ (i \neq j),\ s = 1, ..., 6,\ t = 1999, ...2011\ ,\end{aligned}$$

where d represents outward stocks when the dependent variable describes export flows, and d represents inward stocks when the dependent variable describes import flows. Unfortunately, FDI information is not disaggregated at the sector level, so we can capture only bilateral relations between countries. Focusing on exports we control for the logarithm of the investment that country i holds in country j at the end of year t ($\ln FDI^{outward}_{ijt}$). When treating imports, we control for the logarithm of investment that country j holds in country i at the end of year t ($\ln FDI^{inward}_{ijt}$).[46]

The results are presented as follows: Table 2.9 considers all sectors jointly, column 1 for exports and column 2 for imports; Table 2.10 reports the exports by sector; and Table 2.11 reports the imports by sector. In all cases, the coefficient $\widehat{\beta}_3$ is not statistically different from zero. Though FDI^{d}_{ijt} is a time-variant variable, its variability seems to be captured by the country-pair fixed effects. We can also see that after the inclusion of this new regressor, the significance of our variables of interest does not change; in only a few of cases the magnitudes of the coefficients slightly decrease. Hence, we can declare that our results are robust to the inclusion of bilateral FDI data.

[46] Following the strategy implemented in Sections 2.5 and 2.6, we address the zero values problem by adding one unit to the trade flow data before taking logarithms (Chen, 2004), while missing values are excluded from the sample (Harris et al., 2012).

Table 2.9: EU-10 trade flows: extension 2

Estimator:	FE	
	(1)	(2)
Dependent variable:	ln(exports$_{ijst}$+1)	ln(imports$_{ijst}$+1)
lnGDP$_{it}$	0.9163***	0.9966***
	(0.1156)	(0.1495)
lnGDP$_{jt}$	0.6798***	0.2126***
	(0.0672)	(0.0792)
lnFDI$_{ijt}^{outward}$	-0.0003	
	(0.0013)	
lnFDI$_{ijt}^{inward}$		-0.0036
		(0.0027)
eu15$_{af04}$	0.1368***	0.5910***
	(0.0502)	(0.0545)
eu10$_{af04}$	0.1640**	1.0603***
	(0.0642)	(0.0963)
fsu$_{af04}$	0.0571	0.3640***
	(0.0893)	(0.0899)
constant	-26.0340***	-16.2332***
	(3.2588)	(4.0537)
Observations	59098	52264
Number of groups	8911	7980
R-sq(overall)	0.3689	0.1949
R-sq(between)	0.4053	0.1760
R-sq(within)	0.1462	0.0503
F	175.1592	80.4853
pvalue	0.0000	0.0000

Standard errors clustered by reporter–partner are in parenthesis.

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

Fixed-effects dimension: reporter–partner-sector (λ_{ijs}).

All regressions include time DVs ($af04$ and θ_t).

Table 2.10: EU-10 export flows by sector: extension 2

	Dependent variable: ln(exports$_{ijt}$+1) - Estimator: FE					
	Chemicals	F&B	Machinery	Manufbymat	Othermanuf	Rawmat
lnGDP$_{it}$	0.7045***	0.5391*	1.8612***	0.2763	1.1488***	0.3444
	(0.2250)	(0.3041)	(0.1910)	(0.2300)	(0.2279)	(0.3107)
lnGDP$_{jt}$	0.6082***	0.3081*	0.8600***	0.6283***	0.7721***	0.8657***
	(0.1182)	(0.1717)	(0.1144)	(0.1331)	(0.1251)	(0.1562)
lnFDI$_{ijt}^{outward}$	-0.0012	-0.0029	0.0009	0.0041	-0.0016	-0.0009
	(0.0027)	(0.0038)	(0.0026)	(0.0026)	(0.0022)	(0.0031)
eu15$_{af04}$	0.5285***	1.2577***	-0.2270***	0.2234***	-0.3696***	-0.5646***
	(0.1064)	(0.1363)	(0.0839)	(0.0839)	(0.0884)	(0.1273)
eu10$_{af04}$	0.1972	0.7487***	-0.1322	0.5315***	-0.2009**	-0.1014
	(0.1198)	(0.1676)	(0.1041)	(0.1186)	(0.1001)	(0.1650)
fsu$_{af04}$	0.3588***	0.3921**	0.0810	0.4897***	-0.3079**	-0.6566***
	(0.1369)	(0.1914)	(0.1583)	(0.1699)	(0.1537)	(0.2283)
constant	-19.3271***	-7.9886	-52.2712***	-9.0323	-34.3097***	-17.7407**
	(6.2999)	(8.3203)	(5.3130)	(6.6637)	(6.1588)	(7.9633)
Observations	9953	8629	11201	10370	10787	8158
Number of groups	1505	1397	1625	1532	1581	1271
R-sq (overall)	0.3842	0.2718	0.5159	0.4153	0.4149	0.2271
R-sq (between)	0.4259	0.2685	0.5467	0.4723	0.4602	0.2329
R-sq (within)	0.1738	0.1076	0.2696	0.0664	0.1617	0.1942
F	56.6287	30.1166	102.8110	26.4668	55.1624	43.9197
pvalue	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Standard errors clustered by reporter–partner are in parenthesis.

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

Fixed-effects dimension: reporter–partner (μ_{ij}). All regressions include time DVs ($af04$ and θ_t).

Table 2.11: EU-10 import flows by sector: extension 2

	Dependent variable: ln(imports$_{ijt}$+1) - Estimator: FE					
	Chemicals	F&B	Machinery	Manufbymat	Othermanuf	Rawmat
lnGDP$_{it}$	-0.1969	0.9539***	1.5514***	1.2707***	1.4440***	0.6928**
	(0.3067)	(0.2871)	(0.2648)	(0.2552)	(0.2439)	(0.3298)
lnGDP$_{jt}$	-0.1087	0.1251	0.9631***	0.0969	0.2668**	-0.1325
	(0.1816)	(0.1454)	(0.1537)	(0.1538)	(0.1343)	(0.1823)
lnFDI$_{ijt}^{inward}$	-0.0074	-0.0123**	0.0008	-0.0002	-0.0016	-0.0025
	(0.0070)	(0.0059)	(0.0053)	(0.0054)	(0.0046)	(0.0077)
eu15$_{af04}$	0.3729***	1.1906***	0.4140***	0.3281***	0.6340***	0.5127***
	(0.1006)	(0.0974)	(0.0865)	(0.0858)	(0.0801)	(0.1253)
eu10$_{af04}$	1.2898***	1.2713***	0.7497***	0.8750***	0.8979***	1.2129***
	(0.1719)	(0.1528)	(0.1553)	(0.1642)	(0.1346)	(0.1879)
fsu$_{af04}$	0.4543**	0.7359***	-0.2092	0.3435	0.4297**	0.3914*
	(0.2210)	(0.1928)	(0.1827)	(0.2274)	(0.2066)	(0.2175)
constant	20.6052**	-12.5772	-47.7998***	-19.8707***	-28.9018***	-0.5275
	(8.2483)	(7.8966)	(7.8217)	(7.3631)	(6.8984)	(8.8250)
Observations	7273	8965	9079	8798	10010	8139
Number of groups	1092	1352	1425	1348	1508	1255
R-sq (overall)	0.0230	0.1466	0.4203	0.0759	0.1936	0.0237
R-sq (between)	0.1039	0.1349	0.4268	0.0552	0.1545	0.0047
R-sq (within)	0.0862	0.0649	0.0710	0.0488	0.0476	0.0507
F	26.6196	28.6910	30.4926	26.0563	27.2791	19.3360
pvalue	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Standard errors clustered by reporter–partner are in parenthesis.

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

Fixed-effects dimension: reporter–partner (μ_{ij}). All regressions include time DVs ($af04$ and θ_t).

2.8 Conclusions

In this paper, we investigate the trade integration between the members of the EU. Following Anderson and van Wincoop (2003)'s gravity model, we develop our own model by adding some further *ad hoc* feature. Then, we propose an empirical analysis to evaluate the potential effects of the EU-10's trade flows after the fifth EU enlargement in 2004. Our database contains information on bilateral trade flows (exports and imports) between EU-10 countries and 180 partners, divided into 6 sectors during the period 1999–2011. Using the gravity framework and estimating a DID specification, we analyze how joining the EU has affected the intensity and direction of the EU-10's trade flows.

We find three interesting results. First, we find that trade exchanges between the EU-10 and EU-15 intensified after 2004; however, this trade integration was not homogeneous across all sectors. For sectors that include goods of high technological worth (machinery and vehicles and other manufactured articles), the variation of intensity of the EU-10's export flows to the EU-15 decreased with respect to the reference group, the ROW. This result could suggest that the quality of the production of the EU-10 members either does not always meet the preferences of the EU-15 buyers or that the EU-10 producers cannot compete with those of the EU-15, causing the EU-10 to search for new markets. Second, there was export redirection from the FSU to the ROW between 2001 and 2004, meaning that the past strong commercial connections between the EU-10 and the FSU deteriorated before the official entry of the EU-10 into the EU. Finally, we find that EU-10 consumer demand was strongly biased toward EU-10 products, and this pattern was a generalized effect that extended to almost all the sectors.

Therefore, we detect two reasons that could explain the lack of a complete trade integration between the EU-10 and EU-15 countries, one related to competitiveness and the other to preferences. It would be interesting to further investigate the reasons behind the previous arguments to design and implement public policies more suitable for achieving a deeper economic integration within the EU as a whole. It would also be reasonable to think of more suitable industrial policies to be implemented in the EU-10 countries to foster their competitiveness in a few technologically-sensitive sectors jointly with a clear strategy to favor the diffusion of EU-10 products into the rest of Europe, by promoting, for instance, trade facilities that reduce trade cost.

2.9 Appendix

Table 2.12: The EU member states

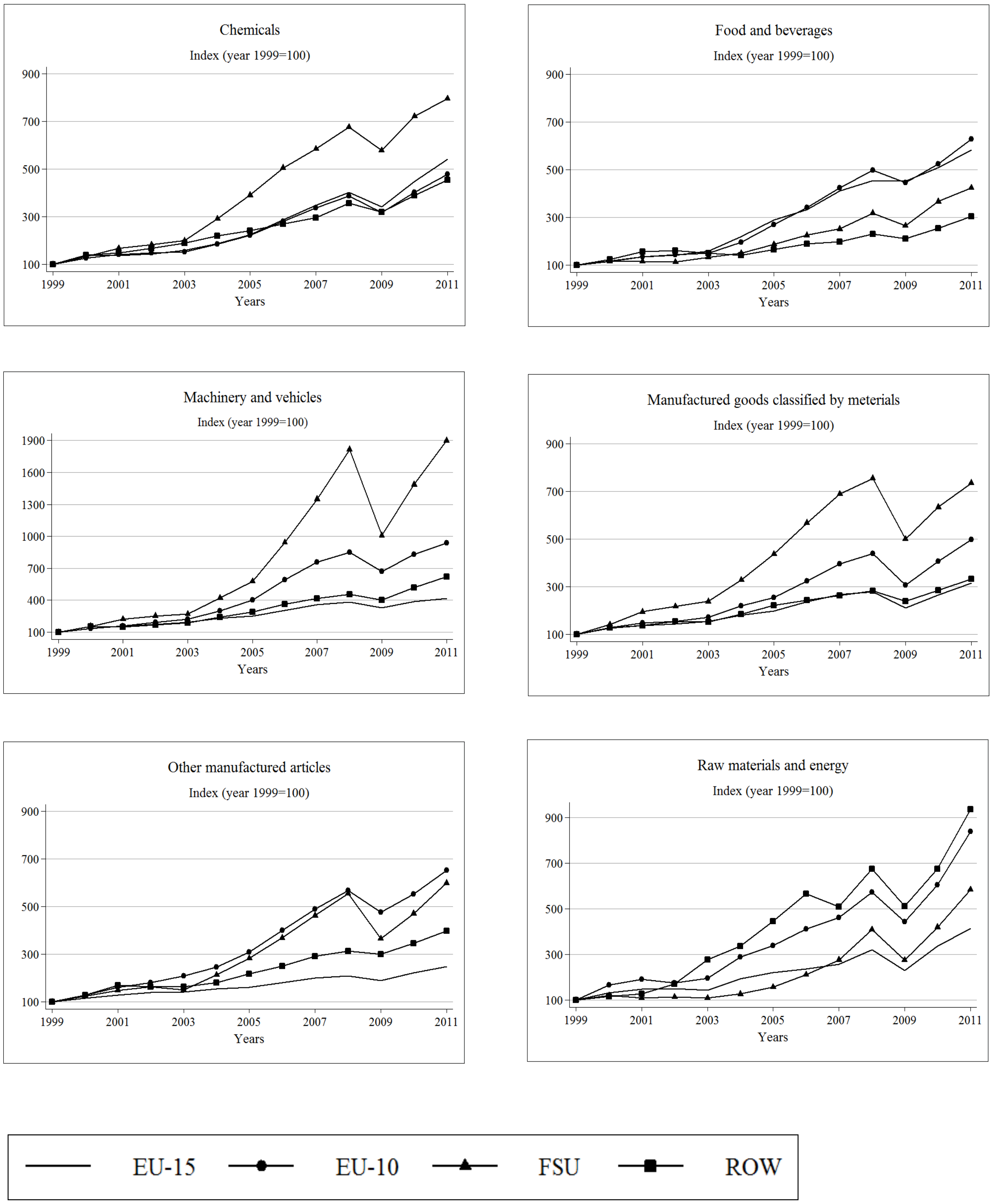
Country	ISO 3-alpha	Year entry	EU subgroup
Austria	AUT	1995	EU-15
Belgium	BEL	1958	EU-15
Cyprus	CYP	2004	EU-10
Czech Republic	CZE	2004	EU-10
Denmark	DNK	1973	EU-15
Estonia	EST	2004	EU-10
Finland	FIN	1995	EU-15
France	FRA	1958	EU-15
Germany	DEU	1958	EU-15
Greece	GRC	1981	EU-15
Hungary	HUN	2004	EU-10
Ireland	IRL	1973	EU-15
Italy	ITA	1958	EU-15
Latvia	LVA	2004	EU-10
Lithuania	LTU	2004	EU-10
Luxembourg	LUX	1958	EU-15
Malta	MLT	2004	EU-10
Netherlands	NLD	1958	EU-15
Poland	POL	2004	EU-10
Portugal	PRT	1986	EU-15
Slovakia	SVK	2004	EU-10
Slovenia	SVN	2004	EU-10
Spain	ESP	1986	EU-15
Sweden	SWE	1995	EU-15
United Kingdom	GBR	1973	EU-15

Table 2.13: List of non-EU countries

Africa (52 countries):
Algeria, Angola, Benin, Bostwana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Rep., Chad, Comoros, Democratic Rep. of Congo, Republic of Congo, Côte d' Ivoire, Djibouti, Egypt, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Lybian Arab Jamahiriya, Madagascar, Malawi, Mali, Mauritania, Mauritius, Morocco, Mozambique, Namibia, Niger, Nigeria, Rwanda, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, South Africa, Sudan, Swaziland, United Rep. of Tanzania, Togo, Tunisia, Uganda, Zambia, Zimbabwe
America (34 countries):
Antigua and Barbuda, Argentina, Bahamas, Barbados, Belize, Bolivia, Brazil, Canada, Chile, Colombia, Costa Rica, Dominica, Dominican Rep., Ecuador, El Salvador, Grenada, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, St.Kitts and Newis, St. Lucia, St. Vicent and the Grenadines, Suriname, Trinidad and Tobago, United States of America, Uruguay, Venezuela
Asia (47 countries):
Afghanistan, *Armenia, *Azerbaijan, Bahrain, Bangladesh, Bhutan, Brunei Darussalam, Cambodia, China, *Georgia, Hong Kong, India, Indonesia, Islamic Rep. of Iran, Iraq, Israel, Japan, Jordan, *Kazakhstan, South Korea, Kuwait, *Kyrgyz Republic, Lao People's Democratic Republic, Lebanon, Malaysia, Maldives, Mongolia, Myanmar, Nepal, Oman, Pakistan, Philippines, Qatar, *Russian Federation, Saudi Arabia, Singapore, Sri Lanka, Syrian Arab Rep., Taiwan, *Tajikistan, Thailand, Timor-Leste, *Turkmenistan, United Arab Emirates, *Uzbekistan, Vietnam, Yemen
Europe (12 countries):
Albania, *Belarus, Bosnia and Herzegovina, Croatia, Iceland, the Former Yugoslav Rep. of Macedonia, *Republic of Moldova, Norway, San Marino, Switzerland, Turkey, *Ukraine
Oceania (10 countries):
Australia, Fiji, Kiribati, New Zealand, Papua New Guinea, Samoa, Solom Islands, Tonga, Tuvalu, Vanuatu

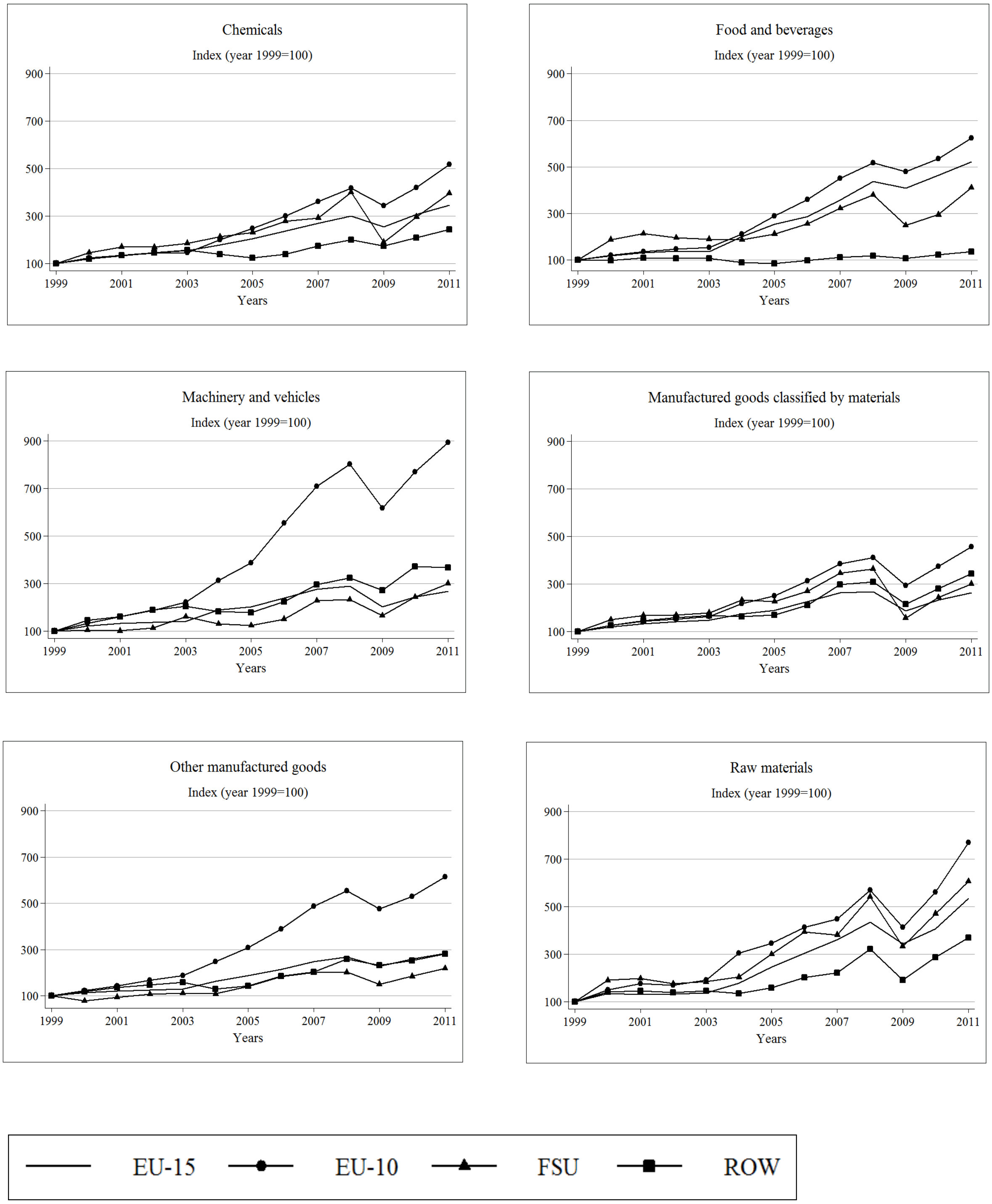
*FSU countries: 12 members (without taking into account Estonia, Latvia, and Lithuania).

Figure 2.4: EU-10 export flows by trading partner and sector



Source: Own elaboration based on data from Eurostat.

Figure 2.5: EU-10 import flows by trading partner and sector



Source: Own elaboration based on data from Eurostat.

Figure 2.6: Trend of the commercial partner DVs over time: EU-10 exports, grup EU-15

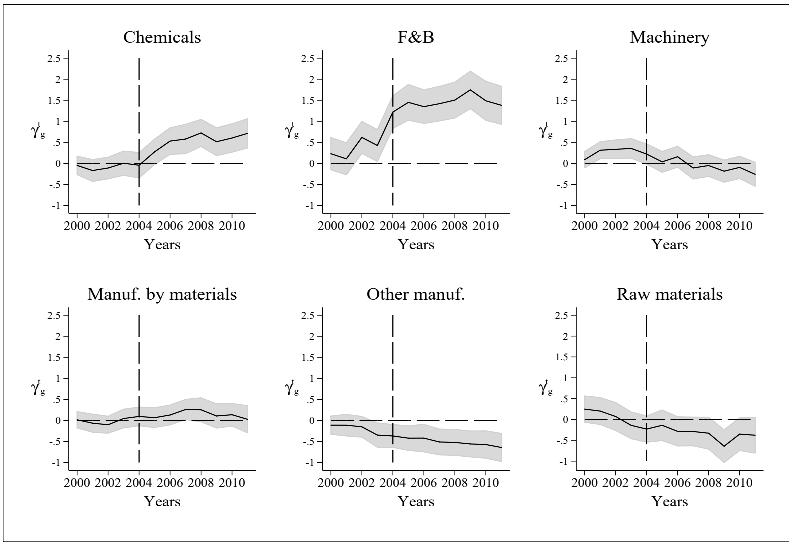


Figure 2.7: Trend of the commercial partner DVs over time: EU-10 exports, group EU-10

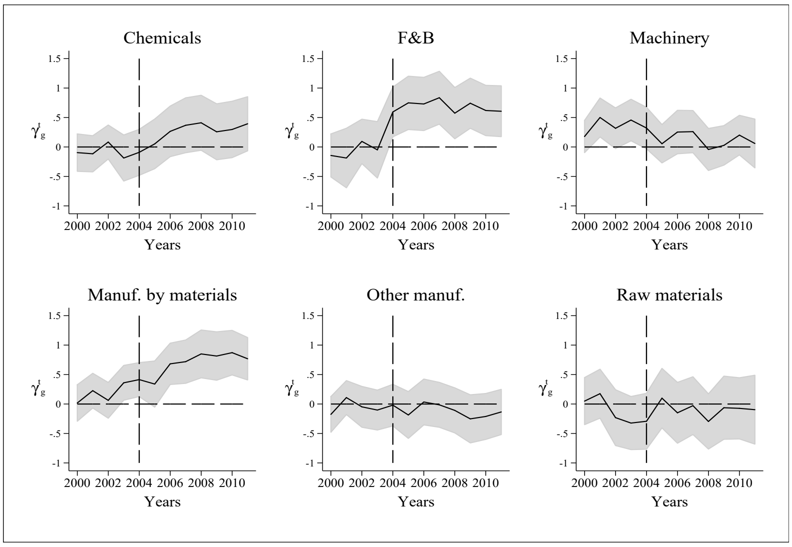


Figure 2.8: Trend of the commercial partner DVs over time: EU-10 exports, group FSU

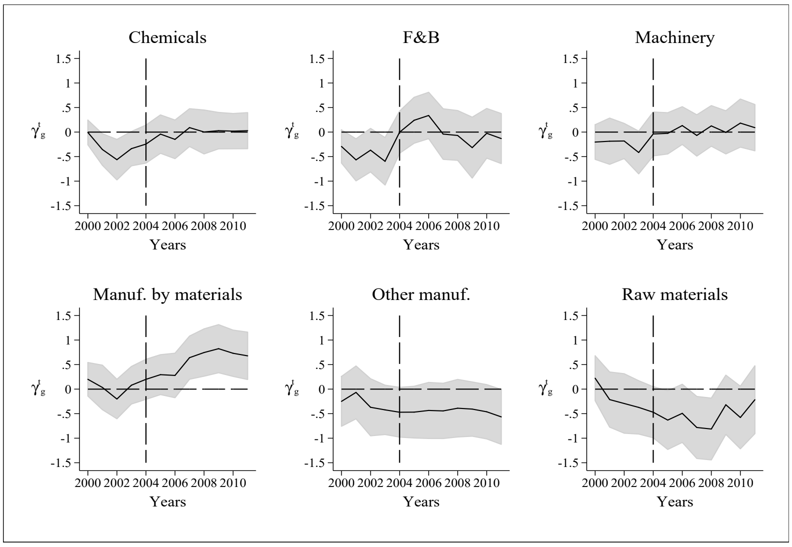


Figure 2.9: Trend of the commercial partner DVs over time: EU-10 imports, group EU-15

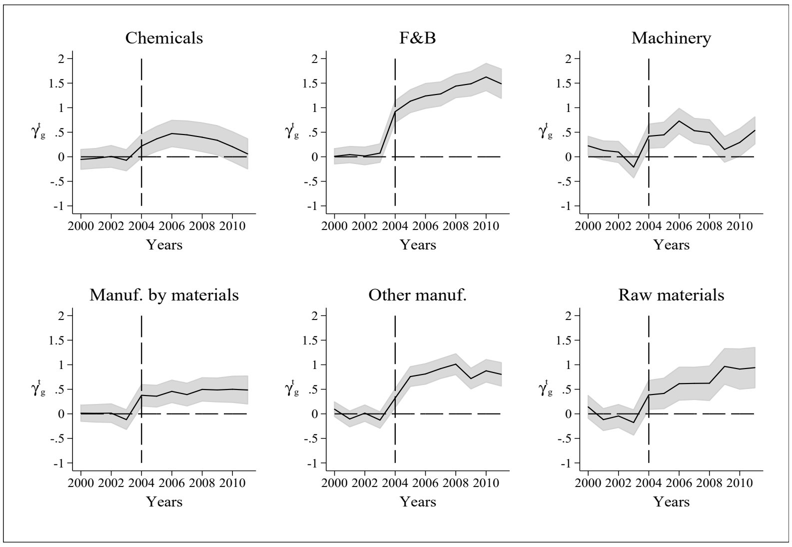


Figure 2.10: Trend of the commercial partner DVs over time: EU-10 imports, group EU-10

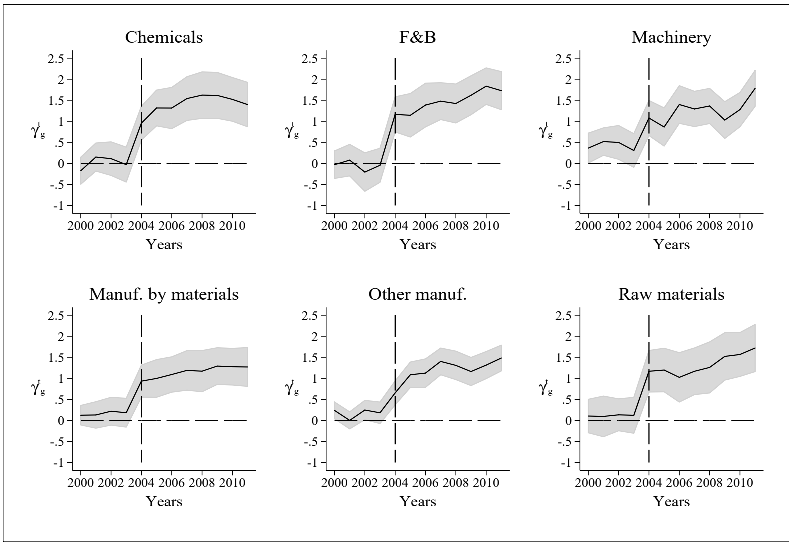
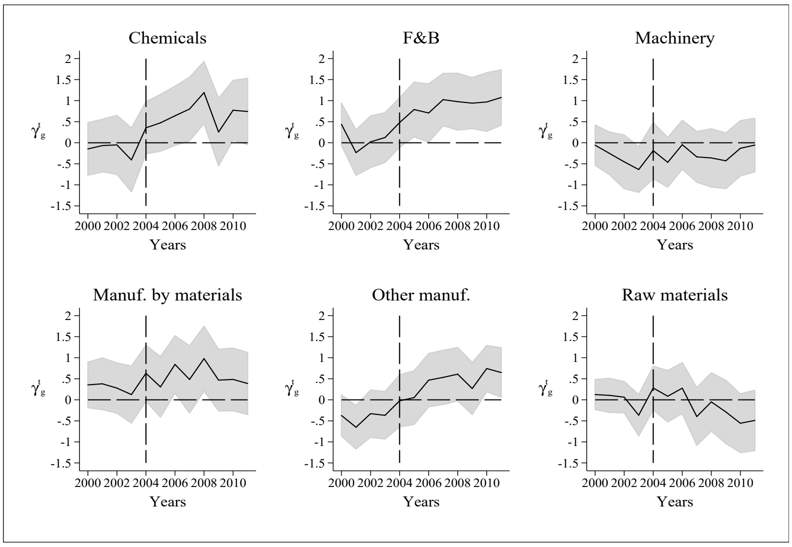


Figure 2.11: Trend of the commercial partner DVs over time: EU-10 imports, group: FSU



Chapter 3

Foreign Direct Investment and Immigration Inflows in Spain[1]

3.1 Introduction

During the second half of the twentieth century, increasingly pervasive globalization induced remarkably stronger flows of foreign direct investment (FDI) around the world. From 1970 to 1997, the worldwide nominal FDI grew by nearly 31 percent, whereas nominal gross domestic product (GDP) and international trade flows, measured by worldwide nominal import flows, increased by 7 percent and 12 percent, respectively (Brakman et al., 2011 p. 324).

Such features were particularly crucial for EU countries, where the removal of barriers to trade exchanges and capital mobility strongly affected trade patterns and volumes, as well as FDI flows, both in and out the European Union (EU; see for instance, Barrell and Pain, 1999; Lafourcade and Paluzie, 2011).[2]

The activity of foreign multinational enterprises (MNEs) is thought to pose advantages for a host economy, mostly in terms of economic performance, productivity, and job creation (Barba Navaretti, 2004; Barrell and Pain, 1999; Crozet et al., 2004).[3] Given these anticipated benefits, governments are often willing to implement costly public policies, including tax holidays and tax waivers, in order to attract foreign investors (Amiti and Javorcik, 2008; Blalock and Gertler, 2008; Haskel et al., 2002). However, if FDI is ultimately ineffective, then it produces no tangible effects in that economy (Konings, 2001). Indeed, an important critique versus FDI as

[1]This chapter benefits from comments and feedbacks from the participants to the Applied Lunch Seminar at UAB (Barcelona, 2015), the Doctoral Workshop of the Department of Applied Economics at the UAB (Barcelona, 2015), the XII Jornadas sobre Integración Económica (Valencia, 2015), and the XREPP's Doctoral Day (Barcelona, 2015).

[2]In 1980, the EU-15 inward FDI represented 5.3 percent of the EU-15 GDP, yet represented 22.2 percent in 1999 (Ekholm and Midelfart-Knarvic, 2004, p. 146), an increase due primary to increased intra-EU FDI (Passerini, 2001).

[3]Per Markusen (2002, p. 5), MNEs are firms that engage in FDI—that is, invest abroad in order to establish a subsidiary or gain control over a foreign firm.

a reliable tool for growth is often its volatility; MNEs are liable to temporary exploit location-specific advantages—above all, favorable fiscal treatment—yet suddenly leave to take advantage of ever better agreement elsewhere (Cebrián de Miguel et al., 2007). As a result, instead of implementing costly public policies, governments should implement specific measures to cultivate an attractive long-term environment for FDI, for example, skill-abundance composition in the local labor market, opportunities for technological partnerships, and selective consumer–provider collaborations.

To clarify the effectiveness of these measures, the chief goal of this study is to analyze the extent to which the entry of new foreign firms in Spain has been influenced by locational determinants and labor market composition. In a sense. we bridge two different strands of literature addressing FDI: one focusing on FDI determinants and the other on relationships within the international movement of capital and people. In this context, Spain is an interesting case for two reasons. First, as a member of the EU, Spain fully enjoys EU membership status and participated in the tremendous increase in FDI inflows. Second, the Spanish labor market is somewhat subject to notable structural problems, including higher unemployment and lower productivity,[4] features that clearly work against the attractiveness of Spain as host country. At the same time, Spain in the 2000s recorded remarkably increased inflows of immigrants, from mostly outside the EU (de la Rica et al., 2014), a phenomenon that exacerbated more the natural imbalances of the local labor market.

The bulk of FDI entry in Spain has been associated with the country's accession to the European Economic Community (EEC) in 1986 (Barrios and Strobl, 2002).[5] During the second half of the 1980s, Spain offered location-advantages to MNEs—most of them related to cheaper-cost production options—and consequently the country became a major FDI receiver in Europe (Ferreiro et al., 1997). In fact, during 1986–1991, Spain received 8.8 percent of all EEC FDI inflows, or the fourth most behind the United Kingdom, France, and the Netherlands, as well as 18.9 percent of total intra-EEC FDI, which was second only to the United Kingdom.[6]

However, since the 1990s, Spain has suffered from two important external reallocation of companies, namely during 1992–1999 and 2000–2007 (Myro and Fernández-Otheo, 2008). The first wave involved the relocation of companies to more developed European countries in search of more sophisticated technology, which accordingly affected primarily high-tech industries such

[4]Of all EU-15 countries during 2004–2014, Spain and Greece had the greatest unemployment rates (Eurostat, 2015a, Table 2). In particular, Spanish unemployment rate was 9.2 percent in 2005 and 24.8 percent in 2012.

[5]Along these lines, some evidence suggests that EEC membership had led a significant increases in FDI for both Spain and Portugal (Baldwin et al., 1996; Ekholm and Midelfart-Knarvic, 2008).

[6]Ferreiro et al. (1997) provide additional statistics.

as manufactures of office machinery and computers, electrical machinery, and medical instruments.[7] By contrast, the second wave entailed the relocation of companies to Central and Eastern European Countries and emerging Asian countries, search of new location opportunities that could reduce production costs. In effect, Spain lost its principal advantages as a venue for FDI after EU enlargements in 2004 and 2007.

But nowadays what are the most important factors attracting FDI in Spain? In line with the existing literature, in this study we provide a quantitative assessment of the importance of agglomeration economies, network forces, and labor market composition in attracting FDI. To implement our empirical strategy, we develop a novel database that merges information from two sources of micro-level data: one for firms and the other for workers. To make both sources comparable, we organize available information by considering 6 home-country groups, 50 provinces, and 22 sectors during 2005–2012. Our approach's most novel feature is the study of the impact of locational determinants and labor market composition upon FDI entry by measuring the number of workplaces opened by foreign investors. As such, our measure of FDI is a real-type proxy, not a nominal one.

We develop a simple econometric analysis by using a fixed-effects (FE) estimator and by defining two econometric specifications. In the first, we aim to capture how specific determinants such as agglomeration economies and networks forces influence the opening of foreign workplaces, whereas in the second we focus on the influence of local labor market composition, defined by the skill and origin of newly hired workers. With the second specification, we also compare determinants that can explain the creation of new domestic workplaces, which we consider as our benchmark, and those that can explain the creation of new foreign workplaces. By these means, we seek to assess the differences and similarities, if any, between domestic and foreign companies in terms of their requirements for hiring employees.

Among our results, agglomeration economies (specifically, localization economies) are relevant to explaining the entry of foreign investors in Spain, as is nationality—that is, the presence of other investors for the same geographical area. This finding suggests that the number of new workplaces opened by investors from a specific home-country group positively depends on investments made one year prior by investors from the same home-country group, but negatively on investments made by foreign investors from other groups.

[7] This relocation arguably stems from Spanish industrial characteristics. Midelfart-Kanarvic et al. (2000) point out the primary differences among industrial characteristics between "northern" and "southern" Europe, the latter of which—Spain concluded—is characterized by lower returns to scale, poorer technology, and less skilled labor.

Unlike the empirical evidence found for other studies (Markusen, 2002; Brakman et al., 2011, chapter 8), our results determine that, in Spain, foreign firms do not demand high-skilled employees, since their employment vacancies are mostly filled by medium-skilled workers. This result poses a clear mismatch between the creation of new foreign vacancies in Spain and its labor market composition On the one hand, although MNEs require medium-skilled employees, the percentage of the Spanish working-age population with secondary education is far below the average reported by the Organisation of Economic and Co-operation Development (OECD).[8] On the other hand, even the percentage of the working-age population with higher (i.e., tertiary) education degrees is close to the OECD average—approximately 32 percent in 2012 (OECD, 2015a)—Spanish graduates face problems with finding positions that match their degrees (OECD, 2015b). Therefore, since a highly-qualified labor force is an important FDI determinant in the long-term, further effort is needed to make Spanish high-skilled workers more attractive to foreign investors.

The rest of the paper is organized as follows. Section 3.2 is a review of the relevant literature, after which Section 3.3 introduces our data and the descriptive statistics. Following an explanation of our econometric specifications in Section 3.4, Section 3.5 discusses our primary results. Lastly, Section 3.6 concludes our paper by presenting few implications for policy.

3.2 Literature Review

Referring to the literature addressing FDI determinants, a seminal contribution comes from Markusen (2002), who presents MNEs activity as part of international trade context and constructs an analytical framework to accommodate that view. In his work, Markusen integrates the ownership, internalization, and location framework (John Dunning,1977; 1981) with firm- and country-specific characteristics for a model that relies on *knowledge capital*, a term encompassing a set of intangible elements such as human capital, patents, blueprints, trademarks, and reputation—to which he refers to investigate FDI determinants and patterns. Knowledge-based assets often have a joint-input or public-good property within the firm, and this characteristic facilitates the internalization of companies. In this context, MNEs tend to more intensively exploit knowledge capital than do domestic firms.

In general, FDI can be classified in two types: horizontal and vertical. Whereas horizontal FDI signifies a firm's replication of its production processes abroad in order to meet the de-

[8] In 2012, 22.28 percent of Spanish population aged 25-64 had secondary education, whereas the OECD average for similar educated people of the same age was 43.88 percent (OECD, 2015a).

mand of the new local market (i.e., market access target), vertical FDI suggests that the firm's production processes are geographically fragmented into several stages as a means to reduce production costs (i.e., comparative advantage target). Despite the frequent difficulty of precisely disentangle these two types of investment,[9] empirical evidence shows that most FDI is horizontal (Markusen, 2002, Markusen and Maskus, 2002).

Several countries have exerted significant effort toward attracting FDI for the positive effects that foreign investors are thought to induce in host economies (Barrios and Strobl, 2002; Crozet et al., 2004). Most of these benefits are achieved via technological spillovers—namely, via technology transfer or labor training.[10] A generalized proposition is that knowledge spillovers generated by MNEs support efficiency and production gains for domestic firms (Blalock and Gertler, 2008; Haskel et al., 2002). However, empirical evidence remains inconclusive, for results are not robust against changes in methodology and datasets (Barba Navaretti, 2004; Barrios and Strobl, 2002; Kemeny, 2010). For instance, Griffith et al. (2003) and Haskel et al. (2002) find evidence of positive spillovers in the United Kingdom, Blalock and Gertler (2008) in Indonesia, and Javorcik (2004) in Lithuania, whereas Konings (2001) finds no evidence in Bulgaria, Poland, or Romania. Other studies point out that these positive externalities are effective only when domestic firms have the appropriate "absorptive capacity" (Barrios and Strobl, 2002; Kemeny, 2010) or belong to research and development intensive sectors (Sembenelli and Siotis, 2008). Interactions between foreign and domestic firms are also crucial for making these positive externalities effective for domestic firms (Barba Navaretti and Venables, 2004b; Javorcik, 2004).

Since empirical results concerning how FDI affects a host economy remain far from being conclusive, an important open question for public policy is whether host governments should endorse costly programs—for example, that subsidize the construction of infrastructure, offer tax holidays, and implement duty exemptions—in order to attract FDI (Amiti and Javorcik, 2008; Blalock and Gertler, 2008; Haskel et al., 2002). Given the lack of consensus in answering this question, case-by-case evaluation seems to be necessary.[11]

Following Ekholm and Midelfart-Knarvic (2004), we divide FDI determinants into three groups.[12] The first group includes industry- and firm-specific characteristics, including level of

[9] For instance, Markusen and Maskus (2001) split average affiliate sales into sales designated to the local foreign market as a proxy of horizontal FDI and sales designated to exports as a proxy of vertical FDI.

[10] The other two important transmission channels are product market and factor market (see Barba Navaretti and Venables, 2004b).

[11] For instance, Ireland has enacted deliberate, successful policy in order to attract FDI, from both the EU and other countries worldwide (Barry, 2004).

[12] See Blonigen (2005) and Ekholm and Midelfart-Knarvic (2004) for a review of literature discussing FDI determinants.

scale economies, costs and benefits of disintegrating stages of production, and firm productivity. Along these lines, it is worth mentioning the contribution of Helpman et al. (2004), who by modeling trade and FDI activity assess the way in which productivity, as a measure of firm heterogeneity, is a truly discriminating feature in explaining how FDI activity can serve foreign markets.

A second group of determinants focuses on characteristics at the country level, such as trade costs, tax differentials, production costs, factor endowments, and market size. Carr et al. (2001) exploit country-specific characteristics (namely, size, size differences, relative endowment differences, trade, and investment costs) and certain interactions among these variables in order to assess the magnitude of their effects upon FDI location decisions. Their results indicate that US outward investment is attracted by more skilled labor-abundant countries. Along the same lines in investigating the relationship between education and the location of multinational affiliates, Shatz (2003) find that US multinational companies seek production locations in populations with high levels of education.

The third group includes other factors that bear important weight in the location of FDI, including regional integration and agglomeration economies. Clearly, reduced internal trade costs associated with regional integration can trigger FDI inflows (Barrell and Pain, 1999; Lafourcade and Paluzie, 2011). At the same time, proximity to other firms could play an important role in determining FDI location during the creation of agglomeration economies or external economies of scale (Basile, 2004; Brakman et al., 2011, chapter 8; Figueiredo et al., 2002; Head et al., 1995).

Yet, though domestic and foreign companies are subject to the same market conditions, their performance often differs sharply (Griffith et al., 2004; Markusen, 2002). Empirical evidence strongly suggests that MNEs perform better than national firms in terms of labor productivity because they are usually larger, invest more in research and development, have larger capital endowments, and hire more skilled labor (Barba Navaretti, 2004). Among these perspectives, by exploiting the principal firms location determinants for both domestic and foreign firms in Portugal, Guimarães et al. (2000) find that foreign location choices depend heavily on agglomeration economies and proximity to major urban areas.

Lastly, another important, more recent strand of empirical literature focuses on the relationship between FDI and migration flows. Most of these studies report a positive link between the international movement of capital and people. For instance, Bhattacharya and Groznik (2008) find that the size of a foreign group from a specific country living in the United States, is positively correlated with US investment in that country. Gao (2003) meanwhile finds that Chinese networks in other countries benefit inward FDI from these countries to China. Other papers by authors such as Buch et al. (2006) and Foad (2012) examine the regional distribution of immigrants and inward FDI stocks in Germany and the United States, respectively. Both authors detect a positive correlation between the stock of FDI and the size of the foreign group from the same country. The reason for this positive relationship relies on the network mechanism; for one, since immigrants create business and social networks that reduce the information barriers for their home-country enterprises, the movement of capital between their home and host countries is expected to flow more easily.[13] Empirical evidence suggests that any positive relationship between FDI and migration is stronger as the level of education among immigrants increases (Foad, 2012; Javorcik et al., 2011; Kugler and Rapoport, 2007). This strand of literature is of particular import for Spain, which since the late 1990s until the beginning of the economic crisis in 2008 has experienced a higher rate of immigrant inflows than other European countries (de la Rica et al., 2014). This massive inflow could have affected the local features of Spanish labor market and, consequently, the inflows of FDI in the country as well. This effect could occur, at least according to Docquier et al. (2014), because immigrants in Spain on average tend to be more educated than natives.

3.3 Data

To investigate the determinants of FDI inflows in terms of locational factors and labor market composition, we create a novel database of information representing 2005–2012, a period that interestingly includes both a cycle of expansion (until 2007) and an economic recession (2008–2012).

3.3.1 Database Structure and Relevant Variables

Our database encompasses information from two data sources: Sistema de Análisis de Balances (SABI, source: Bureau van Dijk (BvD)) and Muestra Continua de Vidas Laborales (MCVL, source: Social-Security). The former provides information concerning firms activity, whereas the latter provides information about the Social Security records of workers.

[13]The idea of the network effects was first exploited in studying the relationship between factors of production and trade flows (Combes et al., 2005; de la Mata and Llano, 2013; Gould, 1994; Head and Ries, 1998; Rauch, 2001). Among the earliest studies, Gould (1994) assesses how immigrants' ties with their home countries can foster bilateral trade between home and host countries—*immigrants' ties* meaning knowledge of home-country markets (e.g., cultural preferences and business opportunities) that contributes to reducing information asymmetries.

Although micro data at the firm and worker levels are available, we cannot merge this information given the lack of a joint identifier. In response, we follow other studies (e.g., Buch et al., 2006) by semi-aggregating our raw data. The level of aggregation that allows us to merge available information is that of origin–province–sector. In particular, *origin* refers to one of the seven places from where investors or workers come: Spain and the six home-country groups of Asia–Pacific, Africa, EU-15, Latin America, and North America, and rest of Europe.[14] Meanwhile, *province* refers to the host place, for which we consider 50 Spanish provinces.[15] Lastly, we apply an *ad hoc* classification by sector to render information in the SABI (NACE-93) comparable with the sector classification adopted by the MCVL (NACE-93 and NACE-2009). Our own classification identifies 22 sectors of production, as detailed in Appendix 3.7, Table 3.5.

Our first data source—the SABI—contains information at the firm level extracted from firms' balance sheets available as disclosed by Registro Mercantil.[16] To construct our key variables—that is, the proxy for FDI entry—we adopt the following selection criteria. First, we select firms established in Spain during 2005–2012 with at least two employees; all self-employment enterprises are excluded. Second, we distinguish domestic firms from foreign ones; to qualify as a foreign firm, a firm has to fulfill one of two conditions: have a parent company located abroad or account for a foreign stake holder with at least 10 percent of total capital.[17] If neither of these conditions is fulfilled, then the firm is considered to be domestic. After classifying new firms by country of origin, we identify the number of workplaces—namely, the headquarters and its delegations, if any—created by each firm. We identify new foreign workplaces, which accommodates our definition of FDI inflows, as well as new domestic workplaces.

One advantage of our FDI indicator is its being a real variable unaffected by values-related concerns. According to Markusen (2002), this approach allows MNEs to be considered as real production units in the economy. Following Foad (2012), who considers the number of foreign affiliates in the United States as a proxy for US inward FDI, we similarly focus on the number of workplaces associated with the opening of new establishments by foreign investors.[18]

[14] "North America" includes Canada and the United States, whereas the "rest of Europe" includes all non-EU-15 European countries.

[15] The provinces correspond to the third level of the nomenclature of territorial units for statistics (NUTS 3) according to the Eurostat geographical classification of regions in Spain. Ceuta and Melilla have been excluded.

[16] Though neither random nor stratified, this sample of more than a million Spanish firms has a size that makes it a reliable reference for economic studies at the national level (Duch et al., 2009).

[17] According to the OECD's and the International Monetary Fund (IMF)'s definition, FDI is any investment in which a foreign investor owns at least 10 percent of the ordinary shares of a company and that aims to establish a long-term relationship to influence the firm's management.

[18] We also have information about the number of employees at new firms, which could be used to purport FDI intensity—that is, the number of vacancies created by new foreign firms. The problem is that because the total number of employees at a company is reported by company headquarters, we do not know how these employees are distributed within the delegations, if any.

Our second data source—the MCVL—contains individual, anonymous data extracted and compiled by Spanish Social Security office, which record information regarding individuals living in Spain, including their gender, age, civil status, country of birth, nationality, highest level of education achieved, and employment status.[19]

A person is included in the MCVL in a specific year if he or she fulfills two independent criteria. The first condition is to having a personal identifier number, that is, the *identificador de persona física* (IPF), which in Spain is the Documento Nacional de Identidad for natives and the Número de Identificación de Extranjeros for foreigners. The second condition is being part of the reference population group, which is defined as people with a relationship with the Social Security office in the year of reference, that is, if they are affiliated within any regimen or receive a contributive pension. The final sample is obtained by a simple random sampling method, in which people from the reference population group are selected if their IPF contains some specific digits, yet randomly selected digits.[20] The MCVL not only allows tracking an individual across time (as long as he or she maintains the same IPF and is part of the reference population), but furthermore includes new people registered by automatic devices, as detailed in Social-Security (2015).

To match worker and firm information, we focus on people who have entered the Spanish labor market—namely, new hirees. As before, we aggregate the information for each year at the levels of origin–province–sector, and use as our definition of immigration inflows the number of new foreign workers. We similarly compute new native workers.

In all, our final database contains information about new foreign workplaces and new foreign workers as a proxies for FDI and immigration inflow, respectively, during 2005–2012.[21]

3.3.2 Descriptive Statistics

In this subsection, we provide a descriptive analysis in order to examine the relationship between new workplaces and new hiring.

First, we focus on the average number of vacancies at new firms, which could indirectly be used as a proxy for the size of investment, and distinguish domestic from foreign firms. Figure 3.1 depicts the massive difference in the average size of these two groups of firms; new domestic firms, on average had 8 employees, whereas foreign firms, on average had 48. This result supports the

[19]Social Security office merges their information with the census extracted from the Instituto Nacional de Estadística (INE) and the Personal Income Tax extracted from the Spanish Tax Agency.

[20]For instance, in 2006 the reference population group was 29.3 million people and the sample was 1.17 million people (i.e., approximately 4 percent).

[21]The final database contains additional variables, all explained in the Econometric Specifications section.

stylized fact that MNEs are usually larger than domestic firms (Barba Navaretti and Venables, 2004a). If we compute the average firm size by sector, this difference continues to be larger. In the case of foreign firms, some sectors had more than 80 employees,[22] whereas among domestic firms, only the energy sector (code 800) had more than 20 employees.

Figure 3.2 describes new workplaces by investor origin. Panel A draws its trend over time, and due to high scale differences, we distinguish domestic values in the left-hand y-axis and foreign values on the right-hand y-axis. If we focus on new domestic workplaces, as represented with grey bars, we can observe a decreasing trend: from $228,905$ new workplaces in 2005 to only $63,783$ in 2012. This substantial drop is likely a consequence of the economic crisis. Concerning the origin of foreign investors, foreign workplaces are chiefly created with capital from the EU-15. But the home-country group is better presented in Panel B of Figure 3.2, where foreign workplaces for a specific year are normalized to 100. In this way, we depict the relative importance of each home-country group. In effect, the graph confirms the large relative importance of EU-15 as a foreign investor; excluding 2010, more than 50 percent of foreign workplaces were created with capital from the EU-15. Workplaces created with capital from North America represented around 17 percent of all foreign workplaces created, whereas investments from Asia-Pacific and Africa were insignificant.[23]

Figure 3.3 depicts new hirees by country of birth; Panel A describes the trend over time, whereas Panel B represents the relative distribution by regional cohorts. We observe that new native workers, as represented with grey bars, show a decreasing trend from $257,927$ in 2005 to $185,063$ in 2012, though such was not as strong as in domestic workplaces. As Panel B indicates, immigrants from Latin American countries represented nearly 50 percent of new foreign workers.[24] Other important groups were those of people from Africa and the rest of Europe, which on average represented 16.6 and 17.7 percent of new foreign-born workers, respectively.[25] The relative importance of the other home-country groups was quite constant over time, as the EU-15 represented 10.6 percent, Asia-Pacific 6.1 percent, and North America only 1 percent.

[22] These sectors are food, beverages, and tobacco (code 100), chemical, plastic, and petroleum refinery (code 400), metallurgy and mechanical equipment manufacture (code 500), and hotel (code 1100).

[23] Our results are consistent with the data provided by the United Nations Conference on Trade and Development (UNCTAD, 2014). In 2005, most inward FDI in Spain came from the EU-15 (76.3 percent) and North America (15.08 percent).

[24] de la Rica et al. (2014) point out that in 2008 more than 2 million immigrants entered Spain from Latin America, a number representing nearly 50 percent of the foreign-born working population. For this same year, we find that Latin American workers represented 49.4 percent of new foreign workers, which indicates that our database is consistent with that study.

[25] These results are consistent with other studies of immigration patterns in Spain. For instance, de la Rica et al. (2014) indicate that the most populous immigrant groups in 2011 were from Romania, Morocco, and Ecuador, which represented 12.4, 11.5, and 7 percent of the foreign-born population, respectively.

We describe composition by sector by referring to Figure 3.4 (Panel A for new workplaces and Panel B for new hirees). More than 67 percent of new domestic workplaces were created in only 5 sectors.[26] In the case of foreign workplaces, the concentration is even larger, only 2 sectors accounted for nearly 50 percent of new establishments.[27] Moreover, as Myro and Fernández-Otheo (2008) point out, the presence of foreign capital in some manufacturing activities (e.g., textiles, leather, wood, plastic, mechanical equipment, and electronic machinery) is expected to reduce over time. On the contrary, a starkly different sectoral pattern is depicted for new foreign-born workers. The construction sector (code 900) accounted for 23.5 percent of new hiring of foreigners, but only 1.2 percent of new foreign workplaces. Along similar lines, the hotel sector (code 1100) accounted for 20 percent of new hiring of foreigners, but only 2.4 percent of new foreign workplaces. These results are consistent with those of González and Ortega (2011), who emphasize that immigrant employment in construction, hotel and restaurants, and domestic services rose noticeably in Spain during 1997–2007.

Another important feature of labor force composition is the level of education of newly hired workers. We distinguish three categories of educational attainment (i.e., low, medium, and high skilled) that correspond with primary, secondary, and tertiary education, respectively. Among the total of new hirees registered during the period of study (to be exact, $2,283,996$ new workers) 60.57 percent were low skilled, 28.24 medium skilled, and 11.19 percent high killed. However, according to data provided by the OECD (2015a), 28.51 percent of the working-age population had a college degree in 2005, a percentage that increased to 32.31 percent in 2012. In that sense, our data capture a clear mismatch between the jobs created and the level of education of Spain's working-age population.

We lastly focus on the relationship between level of education attained and economic activity. Table 3.1 reports the relative importance of economic activities in each level of education, showing that most new workers with primary education were hired in the construction (25.5 percent), hotel (17.4 percent), and administration (16.4 percent) sectors. This group roughly coincides with the distribution of new foreign workers among sectors, thereby confirming a strong relation between immigrant and low-skilled labor.[28] Most new workers with secondary education were hired in the wholesale, retail sale, and motor vehicle repairs sector (23.1 percent) and the

[26]These sectors are construction (code 900), wholesale, retail sale, and vehicle motor repair (code 1000), hotel (code 1100), administrative and support activity (code 1900), and other services (code 2200).

[27]These sectors are wholesale, retail sale, and vehicle motor repair (code 1000) and administrative and support activity (code 1900).

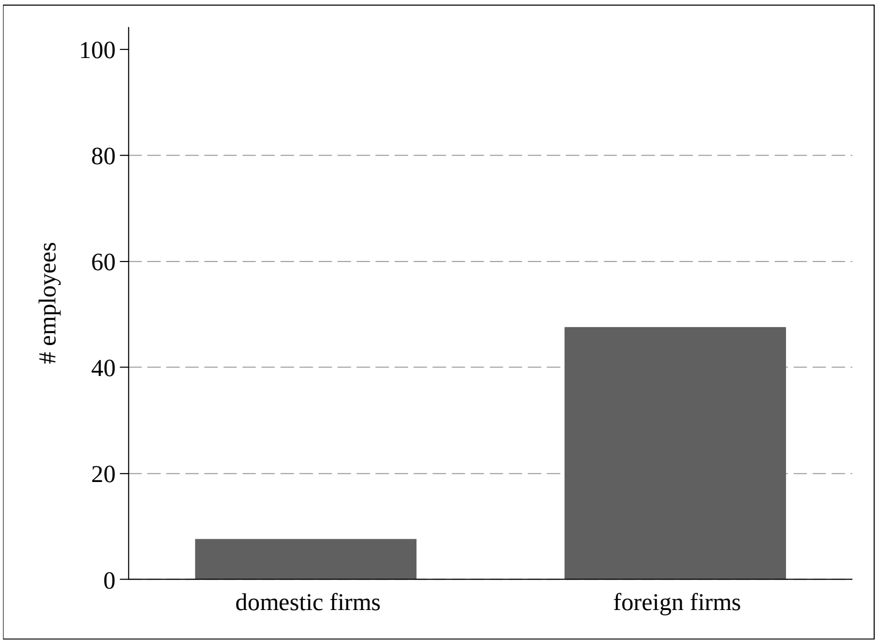
[28]New workers can be classified according to either country of birth or level of education attained, though both classifications cannot be applied at the same time.

administration sector (23.5 percent). At the same time, nearly 40 percent of new workers with tertiary education were hired in services, specifically in education (19.1 percent) and health and leisure (20.2 percent).

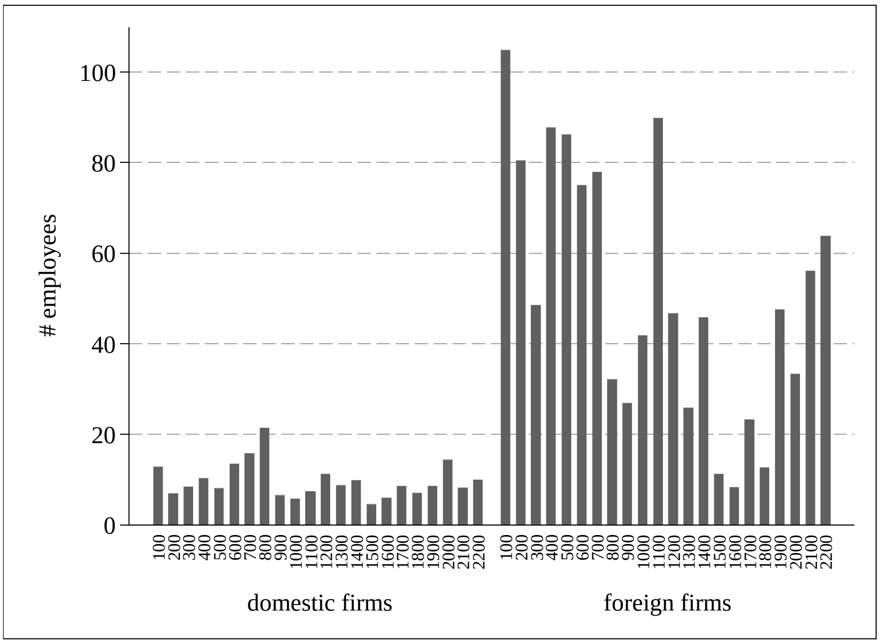
Altogether, our evidence shows that FDI and immigration inflows in Spain follow different patterns. FDI inflows originate primarily from the EU-15 and North America and are concentrated in services. Meanwhile, immigrant inflows originate primarily from Latin America, Africa, and the rest of Europe and are concentrated in the construction, hotel, and administration sectors, all characterized by a preponderance of low-skilled jobs (see Table 3.6 in Appendix 3.7).

Figure 3.1: Average firm size (new firms, 2015–2012)

Panel A: By capital origin



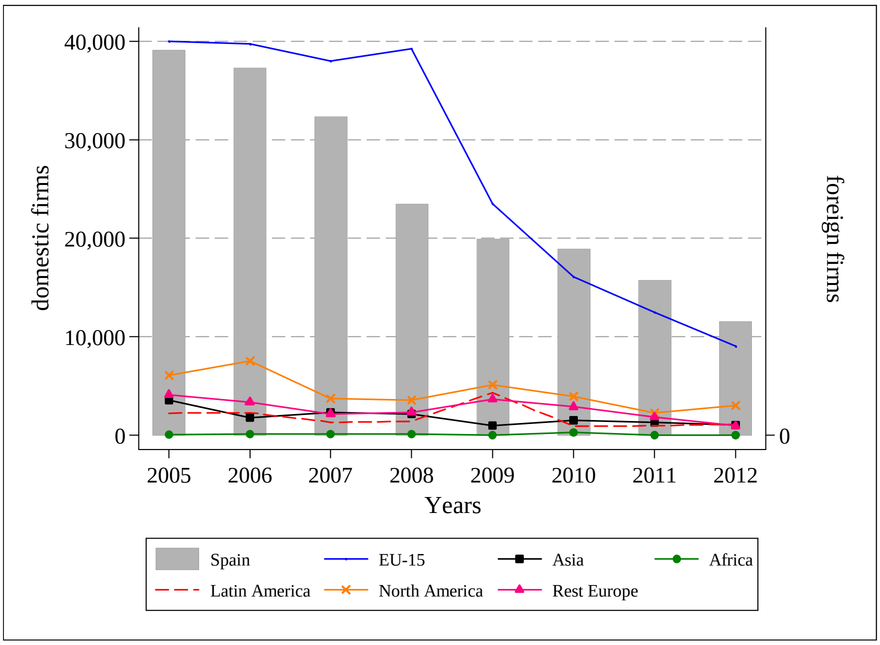
Panel B: By capital origin and sector



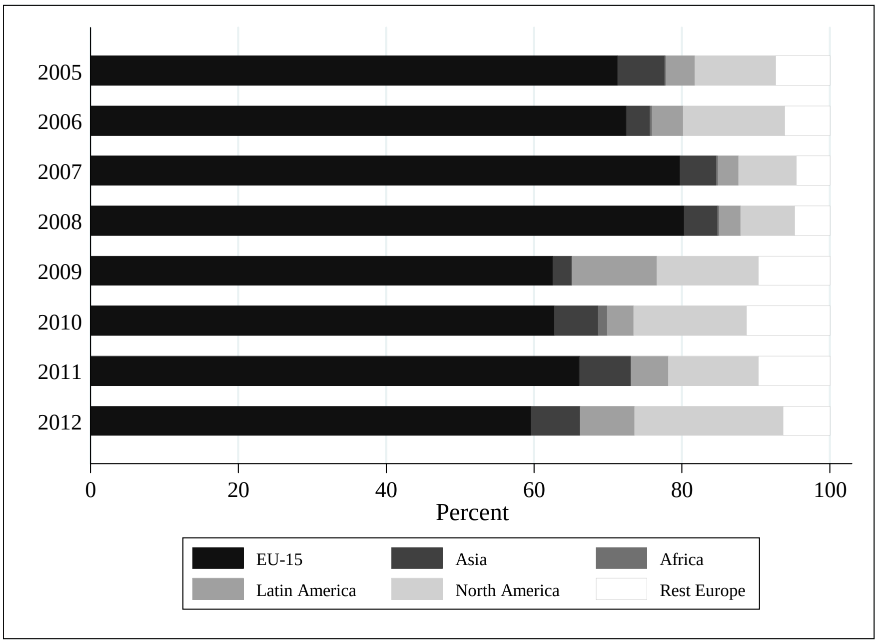
Source: Own elaboration based on data from SABI.

Figure 3.2: New workplaces

Panel A: Trend by origin



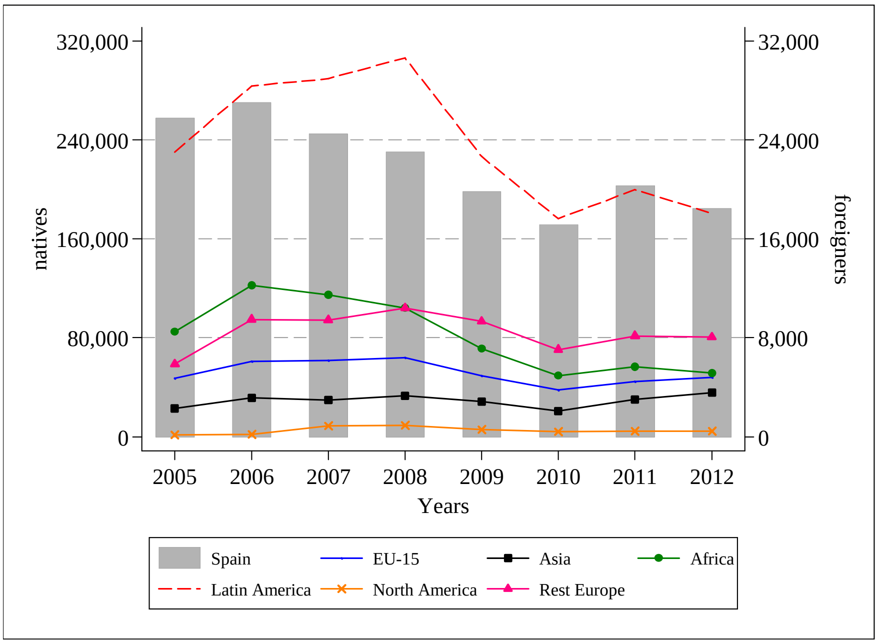
Panel B: Foreign workplaces: distribution by origin



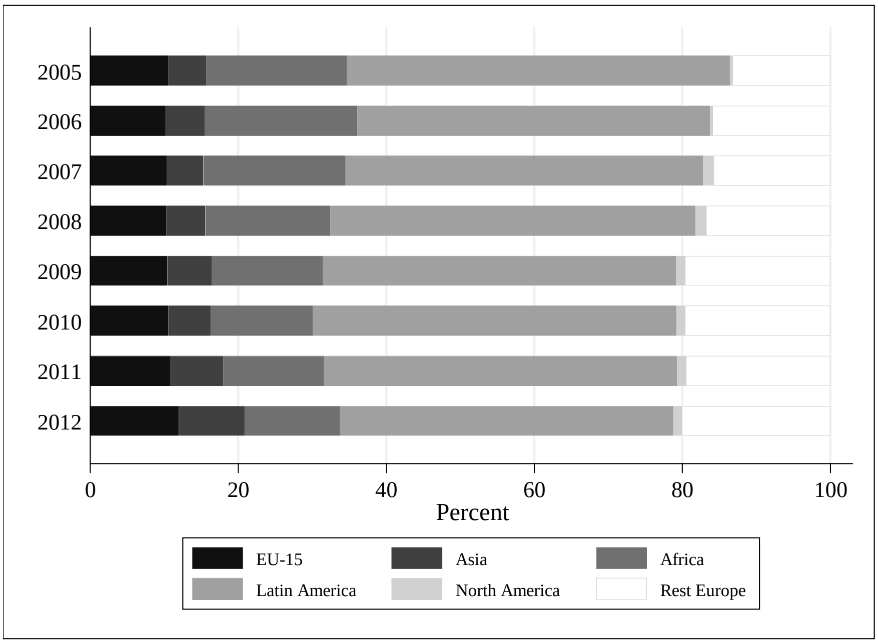
Source: Own elaboration based on data from SABI.

Figure 3.3: New hirees

Panel A: Trend by country of birth

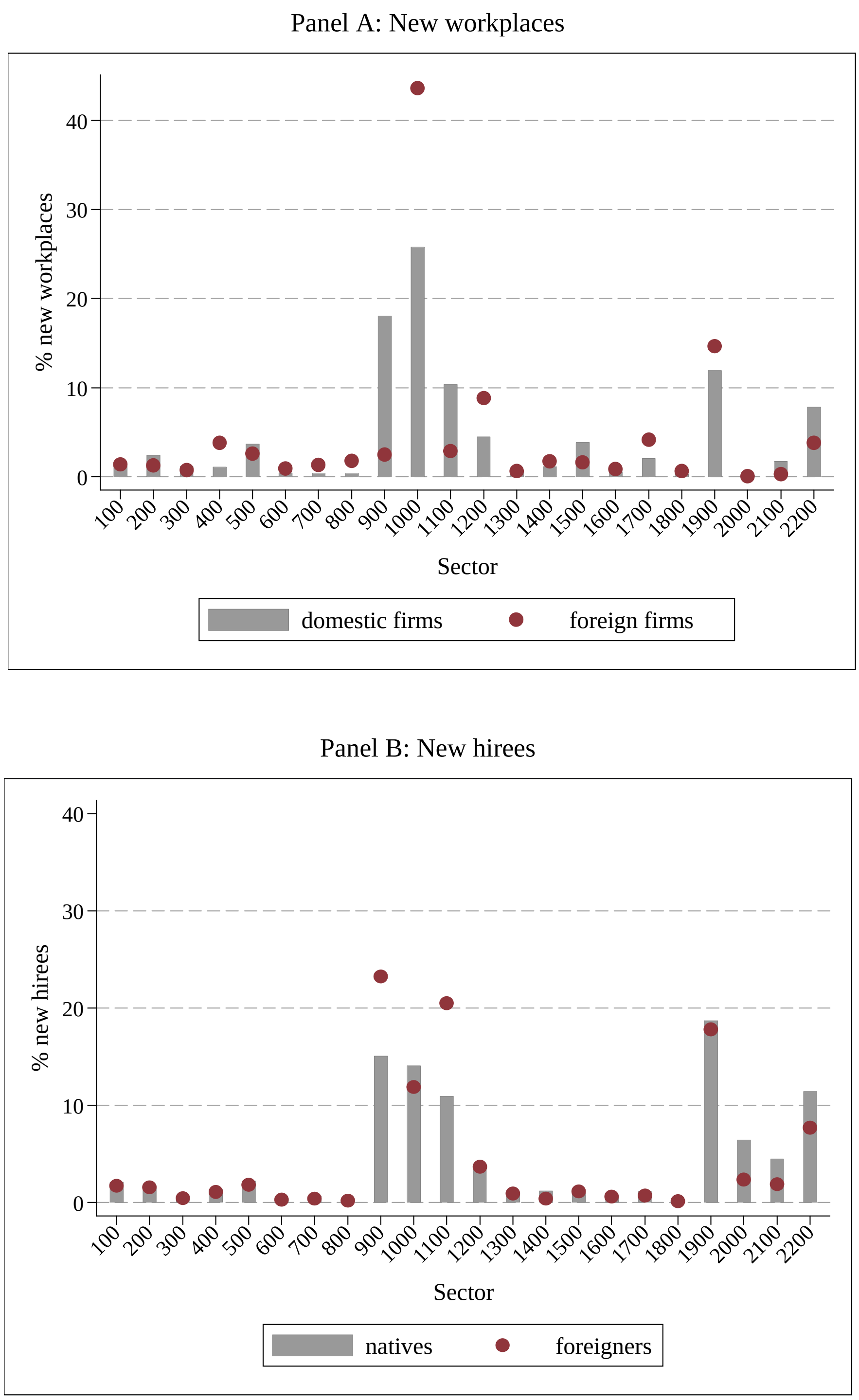


Panel B: Foreign hirees: distribution by country of birth



Source: Own elaboration based on data from MCVL.

Figure 3.4: New workplaces versus new hirees: relative importance by sector, 2005–2012



Source: Own elaboration based on data from SABI (panel A) and MCVL (panel B).

Table 3.1: New hirees by level of education: relative importance by sector, 2005–2012

Code	Sector	Primary	Secondary	Tertiary	Total
100	Food, beverages, and tobacco	2.74	0.97	0.83	2.02
200	Textile, leather, and wood	2.34	0.74	0.44	1.68
300	Paper and publishing	0.52	0.68	0.76	0.59
400	Chemical, plastic and petroleum refinery	1.55	0.83	1.09	1.29
500	Metallurgy and mechanical equipment manufacture	2.95	0.81	1.17	2.15
600	Electrical machinery, computer systems and medical instrument manuf.	0.47	0.25	0.67	0.43
700	Automotive	0.85	0.15	0.63	0.63
800	Energy	0.33	0.18	0.48	0.31
900	Construction	25.50	3.58	5.19	17.05
1000	Wholesale, retail sale, and motor vehicle repairs	10.18	23.09	8.16	13.59
1100	Hotel	17.43	7.24	4.01	13.05
1200	Transport	4.19	3.23	2.18	3.69
1300	Telecommunications	0.30	2.68	0.97	1.05
1400	Financial activity	0.04	2.59	2.59	1.04
1500	Real estate activity	0.58	1.86	1.10	1.00
1600	Renting	0.57	0.72	0.40	0.59
1700	Information technology and computer services	0.24	1.67	4.07	1.07
1800	Research and development	0.03	0.16	1.42	0.22
1900	Administrative and support activity	16.44	23.50	16.62	18.45
2000	Public administration	5.22	5.25	7.93	5.53
2100	Education	0.86	4.67	19.06	3.97
2200	Services (e.g., Health, leisure, sports, and culture)	6.68	15.14	20.20	10.58
	Total	100	100	100	100

Note: The table reports, for each level of education, the relative importance of each sector.

3.4 Econometric Specifications

The core contribution of this study is the investigation of the importance of a group of FDI determinants is Spain. We focus our analysis on agglomeration economies, network forces, and labor market composition, defined by the skill and origin of newly hired workers.

To that end, we implement two different econometric specifications. The former follows the standard specification exploited in empirical studies of agglomeration and network effects, whereas the latter focuses on the role of labor force composition.

The first econometric specification is represented as follows:

$$workp_{opst} = \alpha_1 + \alpha_2 workp_{opst-1} + \alpha_3 workp^{other}_{opst-1} + \alpha_4 workp^{domestic}_{pst-1} + \tag{3.1}$$

$$+\alpha_5 hiring_{opst-1} + \alpha_6 hiring^{other}_{opst-1}\alpha + \alpha_7 hiring^{es}_{pst-1} +$$

$$+\boldsymbol{\beta}\mathbf{V}_{opt-1} + \gamma\mathbf{X}_{pst-1} + \boldsymbol{\delta}\mathbf{Z}_{pt-1} + \mu_{ops} + \theta_t + \varepsilon_{opst} \ ,$$

where $o = 1, ..., 6$ is the home-country group, $p = 1, ..., 50$ is the province, $s = 1, ..., 22$ is the sector, and $t = 2006, ..., 2012$ is the year.

The dependent variable is the number of new workplaces created by home-country group o in province p in sector s in year t. Its values are expected to constitute a linear function of a constant, a group of variables of interest, control variables ($\mathbf{V}_{opt-1}$, $\mathbf{X}_{pst-1}$ and $\mathbf{V}_{pt-1}$), fixed effects (μ_{ops}), year dummies (θ_t), and the error term (ε_{opst}). Following empirical studies (e.g., Alegría, 2006 and Basile, 2004), the explanatory variables are lagged one year because we assume that the act of investing and thereby creating new plants is carried out one year after making the decision to do so.[29]

In Equation (3.1) we introduce a few regressors to capture the impact of agglomeration economies and network forces on FDI inflows. Agglomeration economies merge when companies settle in certain locations, and this concentration triggers benefits in terms of production capacity via spillover effects. Agglomeration is usually represented by the number of companies or the proportion of workers belonging to the same industry that are located in a specific area, as is the case of the so-called localization economies or intra-industry externalities. In that respect, we follow Crozet et al. (2004) by implementing a measure of agglomeration that refers to the nationality of the investor. Our reason for doing so is to confirm whether firms tend to cluster with other firms of same home-country group ($workp_{opst-1}$) instead of other foreign groups ($workp^{other}_{opst-1}$) or domestic firms ($workp^{domestic}_{opst-1}$), if not both.[30] If so, then the coefficient α_2 is expected to be statistically significant and with a magnitude greater than both α_3 and α_4.

Another important group of regressors is that which represents network effects. According to the network effects mechanism, the presence of a sizeable group of immigrants from the same home country should reduce the information costs for home entrepreneurs regarding market considerations in the host economy. Consequently, an increase in capital flows between the two countries is expected. To capture this effect, we introduce new workers from the same home-

[29] According to Alegría (2006), a rational firm opening a plant in location p and in year t, makes its decision based on characteristics that location p had at $t-1$.

[30] For instance, Head et al. (1995) find that Japanese investments in the United States were significantly influenced by previous location decisions of other Japanese firms in the same industry.

country group hired in the same province and sector one year earlier ($hiring_{opst-1}$), new workers from other foreign groups ($hiring^{other}_{opst-1}$), and new native workers ($hiring^{es}_{pst-1}$). Again, if network effects are relevant, then the coefficient α_5 is expected to be statistically significant and greater than α_6 and α_7.

In addition to these proxies, other explanatory variables ($\mathbf{V}_{opt-1}$, $\mathbf{X}_{pst-1}$ and $\mathbf{V}_{pt-1}$) are included in order to control for specific characteristics of local markets that could attract foreign capital.[31] To this end, we select trade openness ($trade\ openness_{opt-1}$) as a proxy for competitiveness (Bloom and Van Reenen, 2007) computed as total trade flows (i.e., sum of exports and imports flows) over GDP.[32] Using the SABI data, we compute a measure of risk or business instability in the host economy as the number of firms (headquarters) that close down in a province–sector combination ($firms\ closing_{pst-1}$).

We also add control variables at the province level. From the INE dataset, we access GDP per capita and the percentage of the working-age population, that is, the population aged 15–64. GDP per capita is transformed into real values ($real\ GDPpc_{pt-1}$) using the Penn World Table (PWT) dataset (Feenstra et al., 2015).[33] For the working-age population, we distinguish natives ($working\text{-}age^{natives}_{pt-1}$) and foreigners ($working\text{-}age^{foreigners}_{pt-1}$).[34] We also include a measure of human capital ($human\ capital_{pt-1}$), represented by the average years of education of the working-age population, and capital stock ($capital\ stock_{pt-1}$).[35]

In addition to the first specification, we propose another to further analyze how labor market composition affects FDI inflows in Spain (Equation (3.2)). Although our dependent variable is again the number of new foreign workplaces, our new explanatory variables here refer to the composition of newly hired employees. We classify members in this group into skill-based categories in accordance to their highest level of education achieved (i.e., $hiring^{highedu}_{pst-1}$ for tertiary education, $hiring^{mededu}_{pst-1}$ for secondary education, and $hiring^{lowedu}_{pst-1}$ for primary education), as well as into the seven regional groups (Spain, EU-15, Latin America, Africa, Asia–Pacific, USA–Canada, and rest of Europe).

The most convenient way to create a proxy of quality of labor is combining the level of

[31] Table 3.7 in Appendix 3.7 provides detailed definitions of all variables.

[32] Information concerning total trade flows is extracted from DataComex (MINECO) and available at the origin–province level. GDP at the province level is extracted from the INE database *Contabilidad Regional de España.*

[33] Specifically, we use the price level of GDP (variable *pl_gdpo*), where the reference is the United States in 2005.

[34] Since the INE includes all EU countries in a group, it is impossible to divide foreigners according to our classification.

[35] Information about human capital is extracted from the Fundación-Bancaja and Ivie database, whereas information regarding capital stock is extracted from the FBBVA and Ivie database.

education attained with the origin of the worker (for instance, new workers from the EU-15 with tertiary education). Nevertheless, we do not dispose that information.[36] To embed this feature while controlling for data limitations, we introduce several interaction terms into our econometric model. Specifically, we host an interaction among the three most representative regional groups of workers in Spain during the period 2005–2012—namely, natives workers, workers from the EU-15 and Latin America—with the three skill categories. The interactions control for the contingent features of the local recruiting process.[37]

Our second econometric specification is thus:

$$\begin{aligned} workp_{opst} = {} & \alpha_1 + \alpha_2 workp_{opst-1} + \alpha_3 workp^{other}_{opst-1} + \alpha_4 workp^{domestic}_{pst-1} + \alpha_h hiring^{highedu}_{pst-1} + \\ & + \alpha_m hiring^{mededu}_{pst-1} + \alpha_l hiring^{lowedu}_{pst-1} + \alpha_{es} hiring^{es}_{pst-1} + \alpha_{eu15} hiring^{eu15}_{pst-1} + \alpha_{lam} hiring^{lam}_{pst-1} + \\ & + \alpha_{h,es} hiring^{highedu}_{pst-1} hiring^{es}_{pst-1} + \alpha_{h,eu15} hiring^{highedu}_{pst-1} hiring^{eu15}_{pst-1} + \alpha_{h,lam} hiring^{highedu}_{pst-1} hiring^{lam}_{pst-1} + \\ & + \alpha_{m,es} hiring^{mededu}_{pst-1} hiring^{es}_{pst-1} + \alpha_{m,eu15} hiring^{mededu}_{pst-1} hiring^{eu15}_{pst-1} + \alpha_{m,lam} hiring^{mededu}_{pst-1} hiring^{lam}_{pst-1} + \\ & + \alpha_{l,es} hiring^{lowedu}_{pst-1} hiring^{es}_{pst-1} + \alpha_{l,eu15} hiring^{lowedu}_{pst-1} hiring^{eu15}_{pst-1} + \alpha_{l,lam} hiring^{lowedu}_{pst-1} hiring^{lam}_{pst-1} + \\ & + \boldsymbol{\beta} \mathbf{V}_{opt-1} + \gamma \mathbf{X}_{pst-1} + \boldsymbol{\delta} \mathbf{Z}_{pt-1} + \mu_{ops} + \theta_t + \varepsilon_{opst} \ , \end{aligned} \tag{3.2}$$

where $o = 1, ..., 6$ is the home-country group, $p = 1, ..., 50$ is the province, $s = 1, ..., 22$ is the sector, and $t = 2006, ..., 2012$ is the year.

We are also interested in assessing potential differences in the weight of the selected determinants between domestic and foreign entrepreneurs. To that end, we replicate Equation (3.2) by considering two dependent variables: workplaces created by domestic entrepreneurs ($workp^{domestic}_{pst}$) and those created by foreign investors ($workp^{fdi}_{pst}$).[38] Referring to estimations of determinants for the creation of new domestic workplaces, we can gauge whether the business environment has a similar or different impact upon the creation of new foreign workplaces.

The longitudinal structure of our database allows us to control for unobservable time-invariant characteristics, as the home–province–sector level (μ_{ops}) in the case of foreign workplaces by origin-country group ($workp_{opst}$) and at the province–sector level (μ_{ps}) in the cases

[36] Taking into account all possible combinations (3 different skill levels, 50 provinces, and 22 sectors) would generate several zeros in our database.

[37] Modelling with interactions allows the marginal effect of one explanatory variable to depend upon the levels of other explanatory variables (Wooldridge, 2006).

[38] Total foreign workplaces are computed as follows: $workp^{fdi}_{pst} = \sum_{o=1}^{6} workp_{opst}$, where $o = 1, ..., 6$ is the home-country group.

of total foreign workplaces ($workp^{fdi}_{pst}$) and domestic workplaces ($workp^{domestic}_{pst}$). We control for unobservable characteristics using the FE or within-transformation estimator.[39]

3.5 Results

With Equation (3.1), we concentrate on the impact of agglomeration economies and network forces on FDI inflows. Table 3.2 reveals interesting results on this point (column 1 excludes control variables, whereas column 2 includes all control variables).[40]

We find that localization economies matter greatly for the location decisions of foreign investors. The coefficient of the variable $workp_{opst-1}$ is positive and statistically significant, meaning that the existence of available workplaces established during the previous period encourages other investors from the same home-country group to invest in the same location.[41] The existence of domestic workplaces established during the previous period also exerts a positive yet minor effect upon foreign investment. However, the presence of other foreign groups imposes a negative impact, possibly related to the effect of competition among firms (Disdier and Mayer, 2004). As such, we can confirm a tendency among foreign investors to cluster in those province–sector combinations where other companies from the same home-country group were established one year before.

Concerning network effects, our results seem at odds with outcomes found in other empirical studies. We conclude that the presence of new workers from a specific home-country group exerted a negative impact on FDI inflows from their home country to their host place. Yet, this result is unsurprising for our specific setting, since our descriptive analysis revealed that capital and labor inflows in Spain follow different patterns (see Figure 3.4). New foreign workers from other groups also have a negative impact on FDI inflows, whereas the number of newly hired native workers has a positive, yet minor impact.

Referring to the control variables, only two variables are positive and statistically significant: trade openness and the percentage of the native working-age population. Our interpretation is

[39] Our dependent variables take non-negative values, and in the case of new foreign vacancies by origin-country group, there are many of zeros (see Table 3.8 in Appendix 3.7). We sought to overcome this dilemma by running a negative binomial (NB) FE estimator, but the model does not perform well statistically. Although the NB random effects estimator converges, it relies on the strong assumption that time-invariant unobservable characteristics are purely random, that is, uncorrelated with regressors. Accordingly, we prefer to use the panel FE or within-transformation estimator.

[40] To reduce the number of zeros, we exclude province–sector combinations, in which no foreign workplace was established during 2005–2012. Though 6,600 groups are possible (i.e., $6 \cdot 50 \cdot 22 = 6,00$), following this procedure we work with only 2,838 groups.

[41] Specifically, *ceteris paribus*, if the number of new workplaces from home-country group o established in province p and sector s at year $t-1$ increases by 10 units, then the new workplaces from home-country group o in this province–sector at year t increases by 3 units.

fairly straightforward. A dynamic commercial relationship between a given country and a given province encourages investors from that country to establish firms in that province. It could signify competitiveness, for the more a province is open to the external market, the larger the amount of FDI inflows. We observe that the percentage of the native working-age population positively affects the number of new foreign workplaces, meaning that foreign investors have incentives to locate in provinces with a larger availability of workers.[42] Although we identify this quantitative effect, we cannot isolate any effect related to the skill of the workers; the variable referring to human capital does not have any impact on the opening of foreign vacancies.

This last issue is better addressed in Equation (3.2), the results of which are summarized in Table 3.3, where column 1 refers to new hirees by skill attainment, column 2 to new hirees by country of birth, and column 3 to selected interaction effects. In these three columns, agglomeration economies pose a positive, significant effect upon FDI entry, thereby indicating that the results found in Table 3.2 are robust.

Concerning the level of education (column 1), we observe that the hiring of medium-skilled workers benefits FDI inflows, whereas the hiring of high- and low-skilled workers has a negative effect. These results counter those of other empirical studies, according to which most FDI is concentrated in skill- and technology-intensive industries. It seems that MNEs invest in Spain to perform standardized-type activities that require medium-skilled employees instead of high-skilled ones.

When we distinguish the labor force by regional cohorts (column 2), we find that immigrants from North America (i.e., the United States and Canada) generate a positive effect upon FDI inflows, whereas immigrants from the Asia–Pacific region generate a negative effect upon FDI inflows. This latter result is not entirely surprising, since official reports released by Spanish institutions, including Social Security office and the Agencia de Trabajadores Autónomos (ATA), show that a large amount of Asian immigrants in Spain—specifically Chinese ones—are self-employed.[43] These people usually set up their own business and remain disconnected from the international FDI movements.

The interaction terms (column 3) show that new workers from Spain and the EU-15 attract

[42] *Ceteris paribus*, province with 10 percent more native working-age population have on average 0.3 more new foreign workplaces.

[43] At the end of 2004, the number of foreign self-employees in Spain was 85,409, of which 11,112 were Asian (Social-Security, 2004). At the end of 2012, these numbers increased to 85,409 in the case of foreign-born self-employees and to 29,920 in the case of Asian ones (Social-Security, 2012). These figures imply hat the growth rate of Asian self-employed individuals has been far more significant than that of all foreign-born self-employed people, at a rate of 169.26 percent against 64.63 percent. According to the ATA (2014), at the end of 2012, self-employed Chinese-born immigrants in Spain represented 18.47 percent of the total foreign-born self-employed.

Table 3.2: Foreign workplaces by origin (1)

Estimator: Fixed effects Dependent variable: workp_{opst}	(1)	(2)
workp_{opst-1}	0.3353***	0.3280***
	(0.1241)	(0.1216)
$\text{workp}^{other}_{opst-1}$	-0.0186***	-0.0216***
	(0.0062)	(0.0074)
$\text{workp}^{domestic}_{pst-1}$	0.0040***	0.0028**
	(0.0015)	(0.0013)
hiring_{opst-1}	-0.0028***	-0.0028***
	(0.0005)	(0.0005)
$\text{hiring}^{other}_{opst-1}$	-0.0008**	-0.0009**
	(0.0004)	(0.0003)
$\text{hiring}^{es}_{pst-1}$	0.0008***	0.0008***
	(0.0002)	(0.0002)
Controls:		
trade openness_{opt-1}		2.9884***
		(0.8927)
firms closing_{pst-1}		-0.0058
		(0.0038)
real GDPpc_{pt-1}		-1.608e-05
		(0.0000)
% $\text{working-age}^{natives}_{pt-1}$		0.0308*
		(0.0176)
% $\text{working-age}^{foreigners}_{pt-1}$		0.0012
		(0.0149)
human capital_{pt-1}		-0.0731
		(0.0714)
capital stock_{pt-1}		-3.909e-12
		(0.0000)
constant	-0.2256*	-0.9650
	(0.1170)	(1.4411)
Fixed effects:		
origin-province-sector	Yes	Yes
year	Yes	Yes
Observations	19,836	19,836
Number groups	2,838	2,838
R-sq (overall)	0.6059	0.5066
sigma_u	1.1283	1.3237
sigma_e	1.1351	1.1317
rho	0.4970	.5777

Home-country group (o), province (p), sector (s), year (t).
Standard errors clustered by province–sector are in parenthesis.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

foreign capital in locations with large population of new medium-skilled employees, ($\widehat{\alpha}_{m,es} > 0$ and $\widehat{\alpha}_{m,eu15} > 0$), but not in locations with large populations of low-skilled employees ($\widehat{\alpha}_{l,eu15} < 0$). In the case of immigrants from Latin America, a reverse pattern is clear, since these workers attract foreign capital in locations with large populations of new low-skilled employees ($\widehat{\alpha}_{l,lam} > 0$).

Overall, these results point out that foreign capital in Spain targets a specific combination of level of education and country of birth. In particular, native workers and those from the EU-15 with a medium level of education and workers from Latin America with a low level of education positively affect FDI entry.

Table 3.3: Foreign workplaces by origin (2)

Estimator: Fixed effects Dependent variable: workplace_{opst}	(1)	(2)	(3)
workp_{opst-1}	0.3106***	0.3200***	0.2966***
	(0.1173)	(0.1189)	(0.1110)
$\text{workp}^{other}_{opst-1}$	-0.0404***	-0.0310***	-0.0545***
	(0.0129)	(0.0116)	(0.0199)
$\text{workp}^{domestic}_{pst-1}$	0.0042***	0.0037**	0.0030***
	(0.0013)	(0.0014)	(0.0009)
$\text{hiring}^{highedu}_{pst-1}$	-0.0032***		0.0007
	(0.0009)		(0.0013)
$\text{hiring}^{mededu}_{pst-1}$	0.0032***		0.0006
	(0.0008)		(0.0007)
$\text{hiring}^{lowedu}_{pst-1}$	-0.0008***		0.0000
	(0.0003)		(0.0002)
$\text{hiring}^{es}_{pst-1}$		0.0006***	0.0002
		(0.0002)	(0.0003)
$\text{hiring}^{eu15}_{pst-1}$		0.0061	0.0056
		(0.0047)	(0.0036)
$\text{hiring}^{lam}_{pst-1}$		0.0011	0.0005
		(0.0017)	(0.0013)
$\text{hiring}^{aspa}_{pst-1}$		-0.0092**	
		(0.0040)	
$\text{hiring}^{af}_{pst-1}$		-0.0049	
		(0.0033)	
$\text{hiring}^{usacan}_{pst-1}$		0.0144*	
		(0.0078)	
$\text{hiring}^{reurope}_{pst-1}$		-0.0038	
		(0.0025)	
Controls:			
$\text{trade openness}_{opt-1}$	3.0075***	3.0166***	2.9721***
	(0.8962)	(0.9007)	(0.8860)
$\text{firms closing}_{pst-1}$	-0.0063*	-0.0063*	-0.0048
	(0.0036)	(0.0037)	(0.0030)
real GDPpc_{pt-1}	-1.769e-05	-1.188e-05	-2.683e-05*
	(0.0000)	(0.0000)	(0.0000)
$\text{working-age}^{natives}_{pt-1}$	0.0140	0.0314*	0.0221
	(0.0184)	(0.0166)	(0.0144)
Interactions:			
$\text{hiring}^{highedu}_{pst-1} \times \text{hiring}^{es}_{pst-1}$			-3.341e-07
			(0.0000)
$\text{hiring}^{highedu}_{pst-1} \times \text{hiring}^{eu15}_{pst-1}$			-3.022e-05

Table 3.3: Foreign workplaces by origin (2) (cont.)

Estimator: Fixed effects Dependent variable: workplace$_{opst}$	(1)	(2)	(3)
			(0.0000)
$\text{hiring}^{highedu}_{pst-1} \times \text{hiring}^{lam}_{pst-1}$			-2.863e-06
			(0.0000)
$\text{hiring}^{mededu}_{pst-1} \times \text{hiring}^{es}_{pst-1}$			3.583e-07*
			(0.0000)
$\text{hiring}^{mededu}_{pst-1} \times \text{hiring}^{eu15}_{pst-1}$			3.122e-05**
			(0.0000)
$\text{hiring}^{mededu}_{pst-1} \times \text{hiring}^{lam}_{pst-1}$			-1.241e-06
			(0.0000)
$\text{hiring}^{lowedu}_{pst-1} \times \text{hiring}^{es}_{pst-1}$			-1.339e-07
			(0.0000)
$\text{hiring}^{lowedu}_{pst-1} \times \text{hiring}^{eu15}_{pst-1}$			-1.083e-05***
			(0.0000)
$\text{hiring}^{lowedu}_{pst-1} \times \text{hiring}^{lam}_{pst-1}$			7.258e-07**
			(0.0000)
constant	-2.808e-01	-9.723e-01	-7.886e-01
	(1.3035)	(1.4703)	(1.3647)
Fixed effects:			
origin-province-sector	Yes	Yes	Yes
year	Yes	Yes	Yes
Observations	19,836	19,836	19,836
Number groups	2,838	2,838	2,838
R-sq (overall)	0.5245	0.4764	0.5869
sigma_u	1.2441	1.3638	1.1493
sigma_e	1.1210	1.1283	1.1079
rho	0.5519	0.5937	0.5183

Home-country group (o), province (p), sector (s), year (t).
Standard errors clustered by province–sector are in parenthesis.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Country of birth: Spain (es), EU-15 (eu15), Latin America (lam), Asia–Pacific (aspa), Africa (af), North America (usacan), rest of Europe (reurope).
Other controls: working-age (foreigners), human capital, and capital stock.

We lastly aim to analyze potential differences between domestic and foreign entrepreneurs. Results shown in columns 1–3 of Table 3.5 refer to domestic workplaces ($workp^{domestic}_{pst}$), whereas those in columns 4–6 refer to foreign workplaces ($workp^{fdi}_{pst}$).[44]

Among results regarding these differences, localization economies are clearly much more important for domestic entrepreneurs than foreign investors. Since we here do not consider investor origin, notable effects found in previous estimations do not appear, thereby stressing

[44] In order to reduce the number of zeros, we exclude the province–sector combinations in which no domestic workplace was set up along the period 2005–2012. There are $1,100$ potential groups (i.e., $50 \cdot 22 = 1,100$), but after this procedure we work with 477 groups.

that localization economies that account for the nationality of investors are critical.

Considering how level of education affects new investment (Table 3.5, columns 1 and 4), locations that hire medium-skilled employees attract new entrepreneurs, both domestic and foreign, whereas locations that hire high-skilled employees produce the opposite effect, especially in the case of domestic entrepreneurs.[45] As before, we observe that foreign workplaces are not located in places where low-skilled employees are hired.

Respecting to the origin-country group of new workers (Table 3.5, columns 2 and 5), we detect some remarkable differences. Although domestic and foreign workplaces seem encouraged by the presence of native workers, this effect is stronger in the case of domestic entrepreneurs. Asian immigrants do not encourage the creation of new workplaces, whether domestic or foreign. As previously discussed, workers in this group are usually self-employed and establish their own businesses. By some contrast, African workers trigger the creation of domestic workplaces only, a result likely associated with the rapid expansion of the Spanish construction sector at the beginning of the 2000s. Workers arriving from North America positively affect the creation of foreign workplaces, but negatively affect domestic ones. These workers have specific professional qualifications, suggesting that their immigration flows are likely linked to established working opportunities within specific affiliates in Spain. Overall, foreign entrepreneurs seek an environment with more qualified workers than do domestic entrepreneurs, as generally confirmed by specifications that include the interaction terms (Table 3.5, columns 3 and 6).

Differences in targets between both groups of entrepreneurs take supported from other results as well. Trade openness is positive and statistically significant at 10 percent in the case of domestic workplaces, but not for foreign workplaces.[46] The number of firms that close down negatively affects the number of new domestic workplaces, but not foreign ones. Although the percentage of the native working-age population is important in the creation of workplaces for both types of entrepreneurs, the effect is stronger in the case of domestic entrepreneurs.[47] Furthermore, human capital negatively affects the number of new domestic workplaces. This and the previous result confirm the idea that foreign investors settle in places with a relatively high concentration of human capital.

[45] The difference among coefficients is statistically different from zero.

[46] This coefficient is positive and statistically significant at 1 percent in Tables 3.2 and 3.3. The change in its significance stems from the aggregation of the six home-country groups.

[47] *Ceteris paribus*, if the native working-age population increases by 10 percent, then the number of new domestic workplaces increases by an average of 18 units, though this number drops to 1.8 in the case of foreign workplaces.

Table 3.4: Domestic versus foreign workplaces

Estimator: Fixed effects						
Dependent variable:	*Domestic* workplace ($workp_{pst}^{domestic}$)			*Foreign* workplaces ($workp_{pst}^{fdi}$)		
	(1)	(2)	(3)	(4)	(5)	(6)
$workp_{pst-1}^{domestic}$	0.7262***	0.6475***	0.6650***	0.0252***	0.0223***	0.0182***
	(0.0462)	(0.0419)	(0.0460)	(0.0080)	(0.0086)	(0.0056)
$workp_{pst-1}^{fdi}$	1.0810**	0.9750**	1.0895**	0.1087	0.1651*	0.0241
	(0.4618)	(0.4472)	(0.4485)	(0.0854)	(0.0928)	(0.0743)
$hiring_{pst-1}^{highedu}$	-0.0604*		-0.1318***	-0.0191***		0.0046
	(0.0359)		(0.0382)	(0.0056)		(0.0078)
$hiring_{pst-1}^{mededu}$	0.0209**		-0.0055	0.0191***		0.0033
	(0.0093)		(0.0167)	(0.0051)		(0.0041)
$hiring_{pst-1}^{lowedu}$	0.0074		-0.0102	-0.0046***		0.0002
	(0.0058)		(0.0128)	(0.0015)		(0.0013)
$hiring_{pst-1}^{es}$		0.0286***	0.0441***		0.0036***	0.0013
		(0.0056)	(0.0127)		(0.0010)	(0.0019)
$hiring_{pst-1}^{eu15}$		-0.0962	-0.0793		0.0362	0.0326
		(0.1828)	(0.2174)		(0.0284)	(0.0217)
$hiring_{pst-1}^{lam}$		-0.0187	-0.0976***		0.0064	0.0029
		(0.0216)	(0.0264)		(0.0100)	(0.0080)
$hiring_{pst-1}^{aspa}$		-0.3363**			-0.0550**	
		(0.1528)			(0.0239)	
$hiring_{pst-1}^{af}$		0.1060*			-0.0290	
		(0.0550)			(0.0200)	
$hiring_{pst-1}^{usacan}$		-1.7390***			0.0876*	
		(0.5549)			(0.0471)	
$hiring_{pst-1}^{reurope}$		-0.1635**			-0.0233	
		(0.0639)			(0.0152)	
Controls:						
trade openness$_{pt-1}$	2.6276	3.1861	3.9690*	0.0560	0.0825	-0.0983
	(2.3713)	(2.2397)	(2.2458)	(0.3781)	(0.4102)	(0.3584)
firms closing$_{pst-1}$	-0.1892***	-0.2167***	-0.2048***	-0.0376*	-0.0375*	-0.0285
	(0.0685)	(0.0704)	(0.0724)	(0.0219)	(0.0223)	(0.0180)
working-age$_{pt-1}^{natives}$	1.6971***	1.5895***	1.8112***	0.1272	0.2276**	0.1765**
	(0.5022)	(0.5087)	(0.5058)	(0.1070)	(0.0993)	(0.0833)
human capital$_{pt-1}$	-5.4912*	-4.9498*	-4.5150*	-0.5018	-0.1964	-0.2998
	(2.8893)	(2.8178)	(2.5519)	(0.3792)	(0.4651)	(0.3512)
Interactions:						
$hiring_{pst-1}^{highedu} \times hiring_{pst-1}^{es}$			-7.902e-06			-2.015e-06
			(0.0000)			(0.0000)
$hiring_{pst-1}^{highedu} \times hiring_{pst-1}^{eu15}$			1.137e-03			-1.824e-04
			(0.0012)			(0.0002)
$hiring_{pst-1}^{highedu} \times hiring_{pst-1}^{lam}$			-5.448e-05			-1.715e-05
			(0.0001)			(0.0000)
$hiring_{pst-1}^{mededu} \times hiring_{pst-1}^{es}$			2.903e-07			2.150e-06*
			(0.0000)			(0.0000)
$hiring_{pst-1}^{mededu} \times hiring_{pst-1}^{eu15}$			2.574e-05			1.879e-04**
			(0.0005)			(0.0001)
$hiring_{pst-1}^{mededu} \times hiring_{pst-1}^{lam}$			-8.439e-06			-7.466e-06
			(0.0000)			(0.0000)

Table 3.4: Domestic versus foreign workplaces (cont.)

Estimator: Fixed effects						
Dependent variable:	*Domestic* workplace ($workp^{domestic}_{pst}$)			*Foreign* workplaces ($workp^{fdi}_{pst}$)		
	(1)	(2)	(3)	(4)	(5)	(6)
$hiring^{lowedu}_{pst-1} \times hiring^{es}_{pst-1}$			3.580e-06			-8.001e-07
			(0.0000)			(0.0000)
$hiring^{lowedu}_{pst-1} \times hiring^{eu15}_{pst-1}$			-2.877e-04			-6.496e-05***
			(0.0002)			(0.0000)
$hiring^{lowedu}_{pst-1} \times hiring^{lam}_{pst-1}$			2.702e-05***			4.323e-06*
			(0.0000)			(0.0000)
constant	-80.6345*	-81.3748*	-140.3253***	-1.3486	-5.7805	-4.2572
	(47.4535)	(45.7610)	(48.8466)	(7.7692)	(8.6680)	(8.1876)
Fixed effects:						
province-sector	Yes	Yes	Yes	Yes	Yes	Yes
year	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,331	3,331	3,331	3,331	3,331	3,331
Number groups	477	477	477	477	477	477
R-sq (overall)	0.9645	0.9387	0.9454	0.4867	0.2523	0.6605
sigma_u	9.8010	19.3504	19.1509	4.6725	5.7887	3.5431
sigma_e	16.2833	15.6348	15.6339	2.6304	2.7402	2.4212
rho	0.2659	0.6050	0.6001	0.7594	0.8169	0.6817

Home-country group (o), province (p), sector (s), year (t).
Standard errors clustered by province–sector are in parenthesis.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Country of birth: Spain (es), EU-15 (eu15), Latin America (lam), Asia–Pacific (aspa), Africa (af), North America (usacan), and rest of Europe (reurope).
Other controls: real GDP pc, working-age (foreigners), and capital stock.

3.6 Conclusion

In this paper, we emphasize the importance of the influence of agglomeration economies, network forces, and labor market composition upon the intensity of foreign workplaces creation in Spain during 2005–2012. To complete our analysis, we elaborate a novel database by aggregating and merging information from two different data sources: the SABI, which contains information about firms, and the MCVL, which contains information about workers.

We find that localization economies are relevant to explaining the entry of new foreign firms, whereas human capital does not play any important role. Accordingly, though unlike other empirical studies, we find that a population of high-skilled workers does not affect the specific setting—namely, foreign firms in Spain do not seek highly qualified employees, and their vacancies are filled by medium-skilled workers. This result indirectly confirms that foreign investors do not privilege quality in the local production environment, but instead pursue better opportunity-costs when investing in Spain.

Our results also generally reveal Spain's problem with incentivizing FDI. On the one hand, MNEs seek medium-skilled-type employees, whereas Spanish workers' education falls below the OECD average of secondary education. In this sense, if one aimed to fulfill current MNEs demand in terms of labor, it would be rather convenient to foster vocational education programs. On the other hand, if Spanish authorities aimed to favor the long-term interest of foreign investors in the domestic economy, it would be advisable to adopt the incentive scheme already consolidated in other countries that relies on making high-skilled local workers a key determinant for attracting FDI.

One weakness of this research is the exploitation of aggregated data. Aggregating information entails the loss of important individual features at the firm and worker levels. A valuable extension would thus be to exploit an employer–employee database that allows us to complete a micro-level analysis and control for firm and worker characteristics. Doing so would yield more precise results concerning the effects of local labor force composition on the intensity of FDI inflows.

Finally, according to our database, foreign investors from the EU-15 represent 60–80 percent of new foreign workplaces in Spain. In that case, it would be interesting to provide a tailored investigation referring to the MNEs with headquarters in EU-15 countries, with the aim to further deepen the understanding of the connection between the determinants of FDI entry in Spain with the evolution of the European integration process.

3.7 Appendix

Table 3.5: Classification of economic activity

Code	Description	CNAE-93 Rev 1					CNAE 2009									
100	Food, beverages, and tobacco	15	16				10	11	12							
200	Textiles, leather, and wood	17	18	19	20	36	13	14	15	16	31	32				
300	Paper and publishing	21	22				17	18	58							
400	Chemical, plastic, and petroleum refinery	23	24	25	26		19	20	21	22	23					
500	Metallurgy and mechanical equipment manufacture	27	28	29			24	25	28							
600	Electrical machinery, computer systems, and medical instrument manufacture	30	31	32	33		26	27								
700	Automotive	34	35				29	30								
800	Energy	40	41				35	36	37	38	39					
900	Construction	45					41	42	43							
1000	Wholesale, retail sale, and vehicle motor repairs	50	51	52			45	46	47							
1100	Hotel	55					55	56								
1200	Transport	60	61	62	63		49	50	51	52						
1300	Telecommunications	64					53	61								
1400	Financial activity	65	66	67			64	65	66							
1500	Real estate activity	70					68									
1600	Renting	71					77									
1700	Information technology and computer services	72					62	63	33							
1800	Research and development	73					72									
1900	Administrative and support activity	74					69	70	71	74	78	82				
2000	Public administration	75					84									
2100	Education	80					85									
2200	Services (e.g., Health, leisure, sports, and culture)	85	90	91	92	93	86	87	88	90	91	92	93	94	95	96

Notes: Codes from 100 to 800 refer to industry, code 900 to constrauction, and codes from 1000 to 2300 to services. The primary sector is excluded.

Table 3.6: New hirees by sector: relative importance by level of education, 2005–2012

Code	Sector	Primary	Secondary	Tertiary	Total
100	Food, beverages, and tobacco	81.88	13.50	4.62	100
200	Textile, leather, and wood	84.61	12.44	2.94	100
300	Paper and publishing	53.03	32.58	14.39	100
400	Chemical, plastic and petroleum refinery	72.49	18.07	9.44	100
500	Metallurgy and mechanical equipment manufacture	83.27	10.64	6.09	100
600	Electrical machinery, computer systems and medical instrument manuf.	66.14	16.37	17.49	100
700	Automotive	81.90	6.87	11.23	100
800	Energy	65.61	16.84	17.55	100
900	Construction	90.67	5.92	3.41	100
1000	Wholesale, retail sale, and motor vehicle repairs	45.38	47.89	6.73	100
1100	Hotel	80.93	15.63	3.44	100
1200	Transport	68.69	24.69	6.62	100
1300	Telecommunications	17.46	72.12	10.42	100
1400	Financial activity	2.04	70.09	27.87	100
1500	Real estate activity	35.26	52.35	12.38	100
1600	Renting	58.21	34.19	7.60	100
1700	Information technology and computer services	13.69	43.83	42.48	100
1800	Research and development	8.40	20.64	70.96	100
1900	Administrative and support activity	54.00	35.90	10.10	100
2000	Public administration	57.17	26.77	16.06	100
2100	Education	13.07	33.17	53.76	100
2200	Services (e.g., Health, leisure, sports, and culture)	38.27	40.33	21.40	100

Note: The table reports, for each sector, the relative importance of each level of education.

Table 3.7: Data definition and sources

Variable	Definition	Source
New workplaces	Number of new workplaces. This variable varies by home country, province, and sector.	SABI (BvD)
FDI inflows	Workplaces created by new firms that fulfill at least one of the following conditions: (i) have a parent company located abroad and/or (ii) account for a foreign stake holder with at least 10 percent of total capital. This variable varies by home country, province, and sector.	SABI (BvD)
New hirees	People entering in the Spanish job market. This variable varies by home country, province, and sector.	MCVL, Spanish Social Security
Immigrant inflows	New workers born in a foreign country. This variable varies by home country, province, and sector.	MCVL, Spanish Social Security
GDP / GDP per capita	The INEbase contains information about the GDP in current euros *Contabilidad Regional de España*. Using the price index (*pl_gdpo*) provided by the PWT, nominal values are converted into real ones. We compute the GDP per capita using information about population, which is also available at the INEbase (*Cifras de Población*). This variable varies by province	INE; PWT (Feennstra et al., 2015)
Trade openness	DataComex provides information about trade flows (exports and imports) by province. For each province, total trade flows is aggregated according to the six different (foreign) origin country-groups. Using the GDP, we calculate an index of trade openness as trade flows over GDP. This variable varies by home country and province.	DataComex (MINECO); INE
Risk	Number of firms that close down. This variable varies by province and sector.	SABI (BvD)
Working age	Data about population are extracted from the INE database *Cifras de Población*. We calculate the working age population as people aged between 15 and 64. We distinguish between natives and foreigners. This variable varies by province.	INE
Human capital	Average years of education of the working-age population. This variable varies by province.	Fundación Bancaja and Ivie
Capital stock	Net capital stock. Using the price index (*pl_gdpo*) provided by the PWT, nominal values are converted into real ones. This variable varies by province.	FBBVA and Ivie; PWT (Feenstra et al., 2015)

Table 3.8: Distribution of new workplaces, 2005–2012 (percentage values)

		0	1	2	3	4	5	6	7	>8
Foreign (by origin)[1]	$workp_{opst}$	91.66	4.03	1.73	0.88	0.43	0.31	0.19	0.15	0.62
Foreign (total)[2]	$workp_{pst}^{fdi}$	65.60	13.78	7.45	3.21	1.80	1.92	0.99	1.11	4.14
Domestic[2]	$workp_{pst}^{domestic}$	6.48	6.39	4.71	5.13	4.38	3.30	2.91	2.85	63.85

Notes:

[1]The table reports the percentage of home–province–sector with 0, 1, 2, ... new workplaces.

[2]The table reports the percentage of province–sector with 0, 1, 2, ... new workplaces.

Chapter 4

Land Specialization in Spain: The effects of the Common Agricultural Policy[1]

4.1 Introduction

The Common Agricultural Policy (CAP) is one of the oldest and most controversial policies of the European Union (EU).[2] The CAP was established in 1962 as one of the pillars of the European integration process, and its main objectives were to ensure adequate food supply and to increase farmers' income.

Although important reasons have been claimed for the regulation of the agricultural sector (Staab, 2013), the CAP system has been criticized since its early years as being considered an inefficient and expensive policy.

In the 1970s and 1980s, the CAP represented around two-thirds of the annual EEC budget (Stead, 2007). Although its relative weight has been falling steadily over time, it still absorbs an important part of the EU budget. In 2011 the CAP cost European taxpayers over 55.6 billion euros, representing 42 percent of the total budget.[3] This amount of financial resources is huge compared with the relative weight of this sector in the composition of the European gross domestic product (GDP). To ease the comparison of the data across time, we focus only on the six founding members of the EEC (henceforth EEC-6).[4] The agricultural sector accounted for 11.5 percent of GDP in 1955 and only 0.1 percent in 2011. In terms of employment, the loss

[1]This chapter benefits from comments and feedbacks from the participants of the Applied Lunch Seminar at UAB (2014), the Seminar of the AQR Group at University of Barcelona (2014), the Doctoral Workshop of the Department of Applied Economics at the UAB (2014), and the XVI Conference of International Economics (2015).

[2]The EU was created in 1957 with the Treaty of Rome. Until 1992 it was officially called the European Economic Community (EEC).

[3]Data retrieved from the EU webpage (http://ec.europa.eu/budget/figures/2011/2011_en.cfm).

[4]The six EEC founding members were Belgium, France, Germany, Italy, Luxembourg, and the Netherlands.

of the relative weight has also been important, decreasing from 21.2 percent (in 1955) to only 2.1 percent (in 2011).[5] Nevertheless, agriculture is still responsible for an important part of the land and water resources used in the EU (OECD, 2012).

Other important criticism of the CAP system is the uneven distribution of its benefits, since the biggest farms (usually owned by rich people or big corporations) accumulate the major share of money transfers. In this respect, Baldwin and Wyplosz (2012, p. 251) provide some data to show the uneven distribution of CAP payments in 2008. Almost half of the EU farmers (48.1 percent) received only 2.1 percent of the total payments, whereas the top 17.6 percent of the recipients received 85.3 percent of the money.[6]

Hence, an open discussion exists about the rationale of continuing to implement this communitarian policy. In this paper, we partially address this issue by focusing on a particular case. We propose an original empirical strategy in order to quantify objectively the effects of the CAP in the Spanish agricultural sector.[7]

Spain is an interesting case because it joined the EEC in 1986, when the CAP was already consolidated. This country experienced an important transformation in its agricultural sector during the 1970s.[8] Thus, the possibility of incorporating the CAP measures could be interpreted as a device to further modernize this sector. In addition, as a recipient of around 13 percent of the total payments, Spain is among the countries that benefit the most from the CAP system.[9]

In the literature, a branch of contributions evaluates the impact that the adoption of an agricultural-oriented policy has on a specific crop or territory (for instance, Donald et al., 2002 and Oñate et al., 2007). This paper, however, contributes to the literature in a more general way: we analyze the extent to which the CAP was able to influence agricultural production in Spain.

In order to perform our analysis, we refer to the seminal contribution by Costinot and Donaldson (2012). Their strategy consists of identifying a potential output that is computed by applying a maximization problem that relies on the Ricardian framework and the concept of opportunity costs. The Ricardian model predicts that countries should produce goods for which they are relatively more productive. The main problem with testing this prediction

[5]If we consider the EU-27, the sector accounted for 1.2 percent of GDP and 5.3 percent of employment in 2011. Sources: EC (2012a) and Zobbe (2001).

[6]Similar results can be found in Jeffery (2003) and the BBC's article "Q&A: Reform of the EU farm policy" (2013).

[7]We are not doing a complete policy evaluation (namely, a cost-benefit analysis).

[8]As a consequence of this important transformation, farm workers lost relative weight in the total Spanish workforce, decreasing from 29.5 percent in 1970 to 19.3 percent in 1980 (EC, 2001b).

[9]Source: BBC's article "Q&A: Reform of the EU farm policy" (2013).

is that relative differences in productivity are not always observable, since not all goods are produced everywhere.[10] However, thanks to the nature of the agricultural sector, it is possible, by using an agronomic model, to identify the productivity of all crops in any part of the world. Their approach implies looking at land heterogeneity (namely, differences in crop productivity) and product prices in order to formulate novel predictions about the most convenient spatial specialization of production.

Furthermore, this strategy holds for this sector because it fulfills the key assumption of the Ricardian model: the sources of technological comparative advantages remain quite stable over time. Of course, agriculture experiences changes in production techniques (such as new chemicals, new machinery, and better technology), but their effectiveness is generally reflected in the medium term, and their impact is less radical than in other sectors.

We rely on the creation of a novel database to account for agricultural production in Spain. This database includes 50 provinces and 25 crops and covers the period 1975–2011. It was built using information from two data sources: the Anuario de Estadística Agraria (AEA) and the Global Agro-ecological Zones (GAEZ) project. The key variable to implement in our empirical exercise is potential productivity. Using an agronomic model and information concerning climate, soil, land cover, and protected areas, the GAEZ project provides information about the potential productivity (kg/ha) of each crop. One of the most important characteristics of the GAEZ project is its level of precision. The Spanish territory is divided into 8,178 grid-cells, and a specific level of potential productivity is associated with each grid-cell.

In the wake of Costinot and Donaldson (2012), the key tool in our strategy is comparing and understanding the similarities and differences between actual and potential agricultural output. In order to assess this issue, we regress the logarithm of actual output over logarithm of the potential output after considering several types of fixed effects that control for unobservable characteristics. The estimated coefficient between these two variables is a proxy to predict how far current production is from potential production, and from there can be interpreted as an efficiency measure of the actual agricultural system.

We obtain two interesting results: the former is associated with the effects of the economic integration process and the implementation of the CAP measures, and the latter with the 2003 CAP reform.

First, we detect an improvement in terms of production efficiency during the second half of the 1980s, namely, after Spain joined the EEC. Unfortunately, our identification strategy does

[10] A good could not be produced in a specific country because its productivity is very low, its selling price is very low, or both.

not allow us to identify which part of this improvement is due to being part of the common market and which part is due to the implementation of the CAP measures.

Second, we detect an efficiency improvement with the Fischler reform (2003). The main element of this reform was to break the link between subsidies and the level of production. This produced a reduction in market distortion, since the farmers' production decisions were based more on market conditions (such as demand, production costs, and prices) rather than on subsidy conditions. Therefore, we provide evidence that the decoupling of money transfers from the level of output had a positive impact on real agricultural output in Spain. We then discuss this result from the perspective of future CAP reforms regarding incentive schemes: future reforms should consider lump-sum money transfers that are not subject to output levels or pecuniary conditions.

Finally, we address our attention to checking the robustness of our results. Our benchmark depends on the value of a few factors: potential productivity, market prices, and land area devoted to agricultural production. In order to check the robustness of our results, we propose three further alternative scenarios. In each of them, we modify one of the essential elements used to determine the efficient scenario. In the first robustness check, we consider another proxy for the potential productivity being provided by the GAEZ project. In the second robustness exercise, we replace Spanish prices with Greek and Portuguese prices in order to test the independence of the results from the level of the nominal component. Finally, we consider a new definition of land endowment devoted to the agricultural production. These tests confirm the consistency of the main results.

The rest of the paper is organized as follows. The next section briefly describes and examines the CAP. Section 4.3 details the data, and Section 4.4 presents the identification strategy. Section 4.5 discusses the main results, and Section 4.6 proposes three robustness checks. Finally, Section 4.7 concludes.

4.2 A Brief Overview of the Common Agricultural Policy

The agricultural sector in Europe was drastically affected by the Second World War since natural resources were damaged and infrastructures were destroyed. In order to recover the prewar standard of production, some European countries met at the beginning of the 1950s to discuss the creation of a common policy for agriculture. Talks took place within the Council of Europe and the Organisation for European Economic Co-operation; the main concerns were to ensure adequate food supply and increase farmers' income (Zobbe, 2001). Nevertheless, because of the strong differences between the European countries, negotiations failed to yield an agreement. One issue was that France and the Netherlands wanted a supranational policy, and the United Kingdom was opposed to such a policy (Zobbe, 2001).

The first agreements concerning agricultural matters took place in the Treaty of Rome (1957). The treaty dedicated some of its chapters to the agricultural sector, but it was in very broad terms. Although some objectives were detailed, the mechanisms to achieve these objectives were not specified until 1962 when the CAP was officially established. The CAP started as a price support policy with the objective of keeping agricultural prices high and stable and near *target* prices, which were set by the Council of Ministers every year.[11]

During the early years, the design of the CAP was strongly influenced by the national agricultural policies of the EEC-6 (Fearne, 1997). Different common market organizations (CMOs) were created to manage regulated products (namely, cereals, sugar, dairy, beef, wine, and olive oil). Each CMO defined a set of rules concerning quality requirements and market price support. Other products such as eggs, poultry meat, pork, fruits, and vegetables were also protected by external producers, but the price interventions were more limited, and products such as seed potatoes were not regulated (Silvis and Lapperre, 2010). Therefore, there were important differences in regulation of different crops during the early years of the CAP. In our empirical exercise, we distinguish between the crops that were regulated from the beginning of the policy and the crops that were historically less regulated in order to analyze whether these differences were reflected in actual agricultural output in Spain.

The first criticisms and worries about the implemented policy soon arose. The European Commission pointed out that price support without any form of structural policy (common to all the EEC members) would never result in the achievement of the important objective of raising farmers' income. Sicco Mansholt, European Commissioner for Agriculture from 1958 to 1972 and principal architect of Europe's farm policy (EC, 2012b), insisted on the need to modernize farming in order to raise its efficiency. One of the problems of the CAP has been the slow implementation of any reform because of the existence of sensitive groups and lobbies opposed to any change.[12]

As expected, one of the consequences of the price support policy was that the production of

[11] For instance, import *levies* were imposed in order to ensure that imports never pushed the internal EEC prices down below a defined threshold (Baldwin and Wyplosz, 2012 and Ritson, 1997).

[12] Paniagua-Mazorra (2001) contrasts three different interest groups' viewpoints concerning the development and implementation of agri-environmental policy in Spain.

agricultural products grew faster than consumption. During the period 1973-1988, the annual production growth of agricultural products was 2 percent, whereas the annual consumption growth was only 0.5 percent (Staab, 2013, p. 215). The EEC started to buy the generated surplus, and part of the surplus was sent to the world market. Consequently, two important problems arose during the 1970s and 1980s. The former was related to the communitarian budget, since an important part of it was used to buy the surplus generated by this sector. The latter was the international pressures from non-EEC countries, above all the net exporter countries, which complained about the protectionist measures (above all export subsidies) implemented by the EEC.

A proposed solution to overcome these problems was to establish production quotas and replace the price support policies with measures referring to income payment compensations. Taking into account this orientation, important changes were introduced by the MacSharry reform (1992) and the Agenda 2000: price reductions,[13] farm income compensations, volume restrictions, and market oriented measures, among others. Nevertheless, these measures were not enough to eliminate the overproduction generated by the sector. On the one hand, the production ceilings were established at a high level, and the fines to those producers that exceeded them were low. On the other hand, direct payments were not completely decoupled from current levels of farm output, which generated incentives to produce more.

The decoupling of direct payments from the level of output was introduced with the Fischler reform in 2003.[14] In order to manage these direct payments, the Council of Ministers established the Single Payment Scheme for farmers and landowners. This system came into force between 2005 and 2007. Each country could decide, among three options, how to distribute the direct payments. Spain applied the historic approach; namely, the amount received for each farmer corresponded to the hectares used to grow and the payment received during a reference period. Direct payments were conditional on some minimal standard requirements concerning food safety, hygiene, environment, animal welfare, and land management (also known as cross-compliance). The aim was to contribute to the development of sustainable agriculture as well as to make the CAP more compatible with social expectations (Jongeneel and Brand, 2010).

In 2007 the CAP matters were divided into two pillars. The first pillar dealt with the policies related to income support and market price support, while the second pillar managed new issues

[13]Intervention prices for cereals dropped by 29 percent, for beef, by 15 percent, and for butter, by 5 percent (Staab, 2013, p. 215).

[14]By 2006, 82 percent of EU direct payments were decoupled from specific output or resource levels (Jongeneel and Brand, 2010, p. 193).

related to rural development, the environment, forests, and fisheries. Although the second pillar gained relative importance over time, the first pillar continued to be the most important in terms of budget resources.[15] Concerning the current crops regulated by the policy, in 2007 the Council of the EU passed regulation on the creation of a new single CMO (Regulation 1234/2007). This new regulation supposed a simplification and homogenization of the policy, moving from 21 separated CMOs to only one CMO.[16]

4.3 Data

We create a novel database for the Spanish agricultural sector by considering 50 provinces and 25 crops over the period 1975-2011. We set 1975 as the starting point of the analysis to avoid the effects of the crisis in traditional agriculture in Spain, which took place during the period 1962-1972 (Barciela et al., 2005).

Our final database contains information about actual and potential production by province, crop, and year. The potential production is calculated by following the numerical exercise proposed by Costinot and Donaldson (2012), in which each parcel of land is specialized in the crop that generates the highest land revenue. This variable indirectly reflects the most efficient way to exploit the agricultural resources, and it represents our benchmark or efficient scenario.

In order to build our database, we extract information from two main data sources: the AEA and the GAEZ project. The former contains Spanish historical information about production and prices of 150 crops, while the latter contains information about the potential productivity of 49 crops. Since we have to merge information from these two data sources, the first important step is the selection of crops. We choose 25 crops that are available in both databases and, according to the AEA classification, are divided into 9 groups (Table 4.1).[17]

The first data source—the AEA—has been published by the Ministerio de Agricultura, Alimentación y Medioambiente (MAGRAMA) since 1904. We extract information about actual production and market prices. The former variable is reported by province, crop, and year, while the latter variable is defined as "the price received by the farmer" and is available only

[15] In 2011 the first pillar represented around 31 percent of the total EU budget, while the second pillar represented around 11 percent. Source: EU webpage (ec.europa.eu/budget/figures/2011/2011_en.cfm).

[16] The single CMO was fully active in 2009 and included the following products: cereals, rice, sugar, dried fodder, seeds, hops, olive oil and table olives, flax and hemp, bananas, live plants and flowers, raw tobacco, beef and veal, milk and dairy products, pig meat, sheep and goat meat, eggs, poultry meat, fruits and vegetables, and wine (Silvis and Lapperre, 2010).

[17] The selected crops were representative of the Spanish agricultural sector during the period of study (detailed statistics are available upon request). In the case of citrus, we had to adopt a specific technique to make both data sources comparable (see Appendix 4.8).

Table 4.1: List of crops

GROUPS	CROPS
cereals	barley, maize, oat, (wetland) rice, rye, sorghum, wheat
legumes	dry bean, chickpea, dry pea
tubers	(white) potato
industrial crops	cotton, soybean, sugar beet, sunflower, tobacco
fodders	alfalfa
vegetables	cabbage, carrot, onion, tomato
citrus	representative crop (orange, mandarin, lemon)
grapes	grape (to transformation)
olives	table olives, olive oil

at the national level.[18] From our second data source—the GAEZ project—we extract the key variable to develop our empirical exercise: potential productivity.[19]

The GAEZ project was created by the Food and Agriculture Organization (FAO) and the International Institute for Applied Systems Analysis (IIASA) in 2002 with the aim of providing data and relevant information to farmers, researchers, and politicians. The last and most complete version was published in May 2012 (IIASA/FAO, 2012a). Relying on an agronomic model and exploiting detailed data about climate, soil, land cover, and protected areas, the GAEZ project delivers information about the potential productivity (kg/ha) of 49 crops in all locations around the world.[20] One of the most important characteristics of the GAEZ project is its level of precision: the earth is divided into more than 9 million grid-cells (4,320 columns and 2,160 rows). In our case, Spain is divided into 8,178 grid-cells.[21]

The computation of potential productivity is achieved by referring to some complex algorithms.[22] Furthermore, for a specific crop, different potential productivity values are computed

[18]From 1975 to 2000, data were entered manually into our database. In a few cases, there is no information about real production; these specific cases are excluded from the analysis. Specifically, the AEA contains missing values for the following crops and years: rice (1982–1984), maize (1990), cotton (1975–1992 and 1995), and sunflower (1975–1995). Nevertheless, when price data are not available, we use different techniques to approximate them (see Appendix 4.8). To have all prices in the same nominal currency (euro/ECU), we use bilateral exchange rates reported by Eurostat.

[19]The GAEZ project does not report information for grape to transformation. This crop represented almost 6 percent of actual production in 2011 (author's computation using the AEA information). Since this is a representative crop for Spanish agriculture, we implement a technique in order to approximate potential productivity for grape to transformation (see Appendix 4.8).

[20]Data from the GAEZ project have been exploited in the economic literature. Nunn and Qian (2011) use information about regional suitability for growing potatoes in countries of the Old World (Africa, Asia, and Europe), Costinot and Donaldson (2012) extract information about the potential production capacity of 17 crops in 55 countries, and Bustos et al. (2015) use data on potential yield for soy and maize in Brazil considering different levels of technology.

[21]In order to extract and manipulate the information that is available in raster files, we use the geographical software *ArcGIS 10* (ESRI, 2011).

[22]First, the best conditions in terms of climate and land resources for each crop are established. Then, taking into account the agro-climatic and agro-edaphic constraints, the expected potential productivity is calculated by

by taking into account different scenarios concerning water supply (rain-fed or irrigation), the input and technology levels (low, intermediate, or high), and the time period (historic levels, average levels, or future time periods).

Specifically, in our exercise we use the variable *potential production capacity for current cultivated land* (PPC), defined as the potential average yield (kg/ha) of the current cultivated land share of each grid-cell.[23] Concerning the different assumptions behind the agronomic model, we approximate them as much as possible to the real characteristics of Spanish agriculture. We choose to adopt the assumptions of irrigation and high technology.[24]

In order to justify the selection of the irrigation assumption, we rely on statistics. Figure 4.1 shows the relative importance of the production techniques concerning water supply in the Spanish agricultural sector. For each crop, total actual production is normalized to 100, and the dark and white shadows indicate which part of this production is obtained as rain-fed and irrigation production, respectively. We appreciate that the production of some crops (citrus, rice, and vegetables) comes mainly from the irrigation production system, while the production of other crops (rye, chickpea, and oat) comes mainly from the dry production system. Figure 4.1 illustrates that the irrigation production system is intensively employed in Spanish agriculture. Overall, irrigation is used in Spain more than in other EU countries because large areas of the country have a semi-arid climate with significant droughts throughout the year (Eurostat, 2015b).

4.4 Identification Strategy

The key tool in our empirical strategy consists of defining a hypothetical scenario in which the land is used in the most efficient way, namely, to maximize its revenues. Then, this scenario is compared with the actual one to get elements for interpreting the real situation in the Spanish agricultural sector, and to assess the potential impact of the CAP policies on actual production.

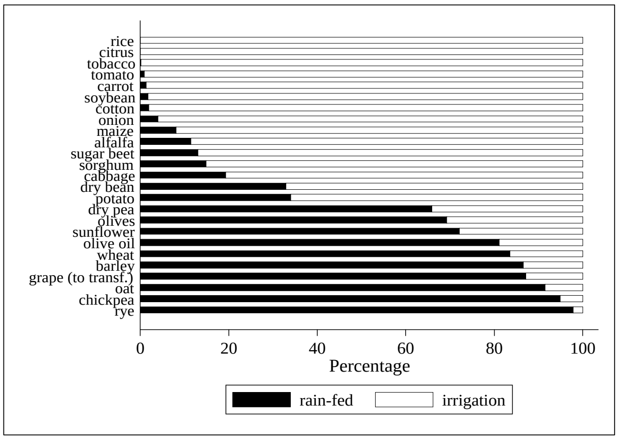
In order to deal with this issue, first we focus on the computation of the benchmark scenario. Then, through a descriptive exercise, we compare our predictions with the real data in order to check the main similarities and differences. Finally, we present the econometric specification.

crop and grid-cell. For a more detailed description, see IIASA/FAO (2012b), Module V.

[23]The PPC in a specific grid-cell is equal to -1 when no share of the parcel of land is being used to agriculture, meaning that this area is not suitable for agricultural production.

[24]High technology assumes a complete market-oriented system where production is fully mechanized and the most advanced techniques are used.

Figure 4.1: Water supply by crop: rain-fed versus irrigation (Spain, 1975–2011)



Source: AEA (MAGRAMA). Calculus: Author.

4.4.1 Potential Production

The benchmark scenario owes a lot to the original exercise developed by Costinot and Donaldson (2012), who calculate predicted output in an economy where the factor of production is allocated according to the Ricardian idea of opportunity costs.[25]

Let's consider an economy with $r = 1, ..., R$ regions, $c = 1, ..., C$ crops, and $f = 1, ..., F$ heterogeneous factor of production, where each f corresponds to a distinct parcel of land. A parcel of land can be used to produce any type of crop, but with different returns. Let us define $A_f^c \geq 0$ as the productivity (kg/ha) of crop c in parcel f.[26] We assume full specialization, namely, a parcel of land f is specialized in only one crop. Each region is endowed with a fixed number of parcels. $\overline{L}_{rf} \geq 0$ denotes the endowment (ha) of parcel f in region r. The total output of crop c in region r is given by

$$Q_r^c = \sum_{f=1}^{F} A_f^c L_{rf}^c \ , \tag{4.1}$$

where L_{rf}^c are the hectares of parcel f allocated to crop c in region r. It represents the land allocation across crops and is the variable we are interested in. We aim at defining the best land allocation across different crops by exploiting all the available land in each region.

We focus on the supply side of the economy, and market prices (p^c) are taken as given. The efficient land allocation across crops is expected to be achieved by the fulfillment of two

[25]Their empirical exercise is based on the theoretical framework presented by Costinot (2009).
[26]$A_f^{c_1} > A_f^{c_2}$ means that the parcel f is more productive in producing crop 1 (c_1) than crop 2 (c_2).

requirements: *(1)* maximizing the value of production in all regions and *(2)* being feasible. Hence, this implies solving the following maximization problem:

$$\max_{L_{rf}^c} \quad \sum_{r=1}^{R}\sum_{c=1}^{C} p^c Q_r^c \tag{4.2}$$
$$st \quad \sum_{c=1}^{C} L_{rf}^c \leq \overline{L}_{rf} \ .$$

The linearity of aggregate output makes its solution relatively easy: each parcel of land f should be employed in the crop that maximizes land revenue (namely, $p^c A_f^c$), independently from the crops that are cultivated in other parcels.[27]

Assuming that the efficient factor allocation is *unique*,[28] the potential (or predicted) output of crop c in region r can be expressed as

$$\left[Q_{rf}^c\right]^* = \sum_{f\in\Omega_r^c} A_f^c \overline{L}_{rf} \ , \tag{4.3}$$

where Ω_r^c is the set of parcels where the crop c is produced, and it is defined as follows:

$$\Omega_r^c = \left[f = 1, ..., F \mid \frac{A_f^c}{A_f^{c'}} > \frac{p^{c'}}{p^c} \text{ if } c \neq c' \right]. \tag{4.4}$$

In other words, crop c is produced only in those parcels in which it brings the maximum land revenue.

Equation (4.4) captures the idea that land heterogeneity plays an important role in determining the patterns of specialization. Costinot and Donaldson (2012) use a sample of 55 countries and 17 crops in the year 1989 to empirically test this idea. They compute predicted output by country and crop. By regressing real output against predicted output (in logarithms), they determine the presence of a positive and significant elasticity, meaning that the Ricardian idea of opportunity cost has significant explanatory power in their analysis. Put differently, this elasticity can be also read as a measure of the efficiency of a current production system: when it is close to one, it implies that the current production system converges near the efficient one, which maximizes the profitability of the available resources. In practical terms, this is also the tool we are exploiting for testing how close we are to the efficient level of production for the agricultural sector in Spain.

[27] Appendix 4.9 provides a further discussion about the assumptions of the model.

[28] This term means that for each parcel of land f, there is only one crop that maximizes its revenue. Costinot (2009) reports that uniqueness is more likely in economies with a large number of regions or a large number of factors. Specifically, in our numerical exercise the efficient land allocation is unique.

In order to define our benchmark scenario, first we have to determine the efficient land allocation by province and year, assuming that the land is used in the most efficient way, namely, to maximize its revenues (Euros/ha). Two variables are essential for calculating land revenues: potential productivity, which is constant over time, and market prices, which are time-variant. The former is a proxy for the GAEZ variable PPC*,* and its value depends on the characteristics of the land (nutrients, slope, etc.), as well as the climate conditions (precipitation, frequency of wet days, mean temperature, temperature range, etc.). The latter variable, market prices, is extracted from the AEA.

It could be interesting to further investigate the behavior of these two variables, since they are crucial for determining the efficient land allocation and, indirectly, potential production. Land heterogeneity across provinces plays a crucial role in determining efficient spatial specialization in agriculture (Costinot and Donaldson, 2012). In the case of Spain, Figure 4.2 illustrates an example of land heterogeneity. The map shows the ratio of potential productivity in wheat relative to potential productivity in olives. Black areas have low productivity in olives and strictly positive productivity in wheat, while in white areas the productivity of wheat is zero. High differences in relative prices could also be crucial in determining the efficient land allocation in a specific year. To this extent, Figure 4.3 presents the real price trend for six different panels.[29] We appreciate that the real price level of dry bean, chickpea, tobacco, and olive oil is considerably higher than the price level for the other crops. These higher prices should increase the profitability of producing these crops, although the agronomic conditions to produce them will not be the most rentable ones.

Once the efficient land allocation is defined, namely, the determination of what to produce in each parcel of land, we consider potential productivity and the parcel dimension to compute potential production by province, crop, and year (Equation (4.3)).[30]

Concerning the parcel dimension, we consider each grid-cell as a different parcel of land f and its total area as the endowment $\overline{L}_{rf} \geq 0$.[31] We assume that all the available land is used for agricultural production except for the urban areas, which are defined according to the shapefile called *Núcleos de Población* published by the Sistema Integrado de Información de Agua (SIA, source: MAGRAMA).[32] Figure 4.4 provides an example using the province of Madrid. The

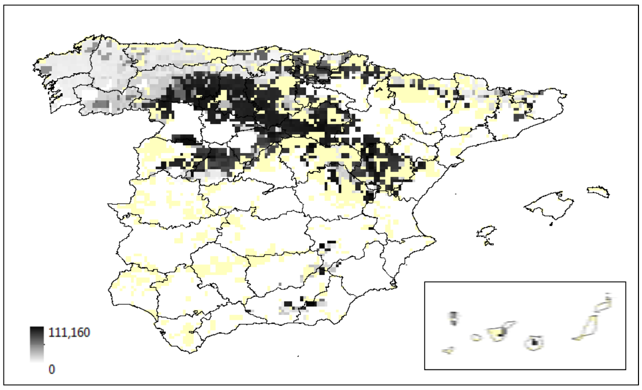
[29] In order to convert the values from nominal to real ones, we use the Penn World Table database 8.0 (Feenstra et al., 2015). Specifically, we use the price level of household (pl_c), where the reference is the United States in 2005.

[30] Appendix 4.10 provides a schematic description of the numerical exercise.

[31] The median grid-cell measures 6,278.91 hectares (ha) or 62.79 square kilometers.

[32] We erase the urban areas from a grid-cells map, and then we match the resulting map with the Spanish political map, provided by the SIA (MAGRAMA).

Figure 4.2: Relative productivity differences: wheat versus olives

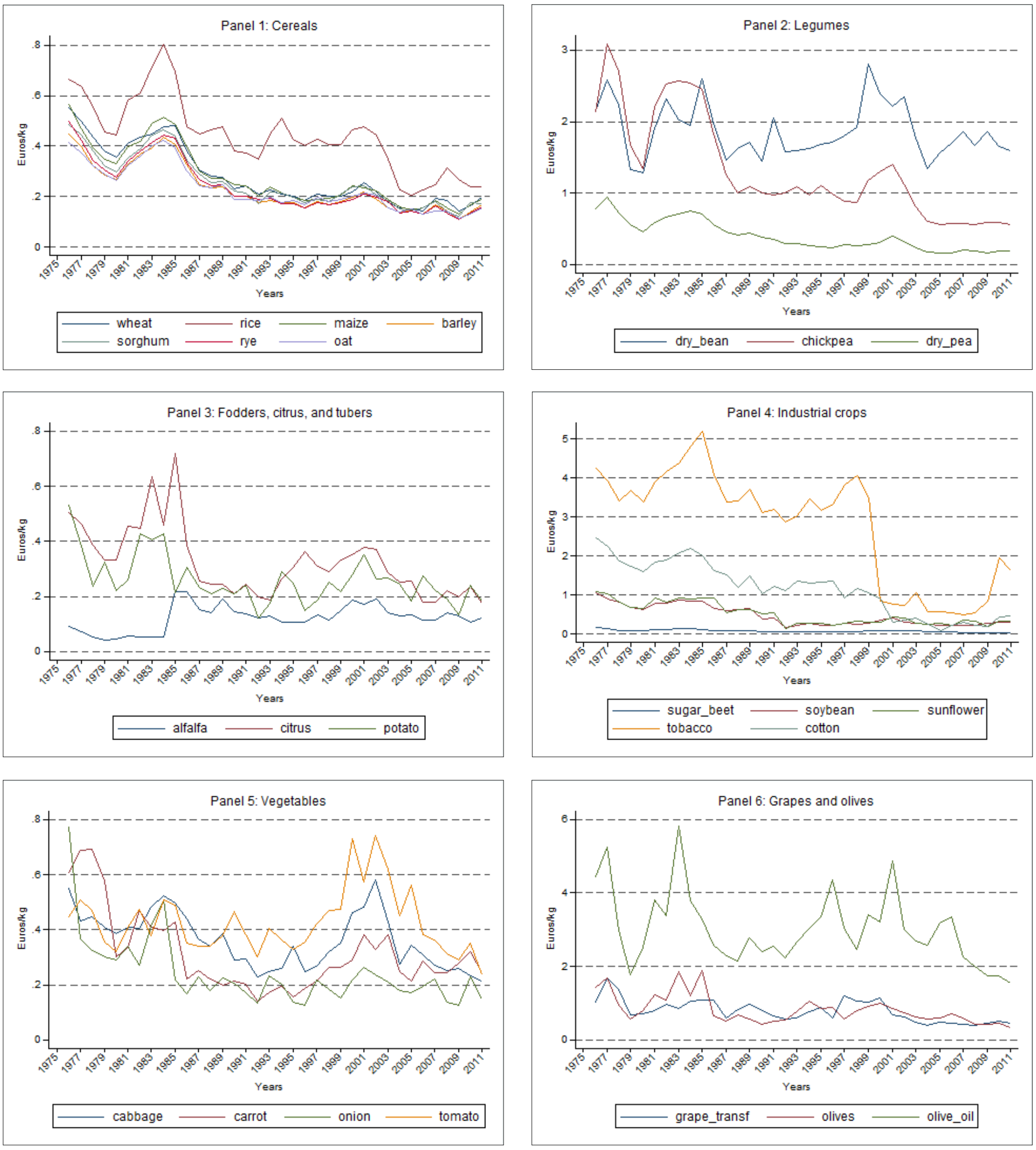


Source: GAEZ project v3.0 (IIASA/FAO, 2012a). Calculus: Author.
Notes:
Variable: PPC (kg/ha). Water supply: irrigation. Technology: high.
Areas shaded yellow are grid-cells in which PPC is equal to -1.

shaded black areas are urban areas not suitable for agricultural production. As we can observe, the same parcel of land f can belong to more than one province. For this reason, it is important to distinguish and quantify the part of the parcel of land that belongs to each province.[33]

[33]To calculate the area of the parcels of land located in a province, we use the geographical software *ArcGIS 10* (ESRI, 2011).

Figure 4.3: Real prices trend (Spain, 1975–2011)



Source: AEA (MAGRAMA, 2013). Calculus: Author.

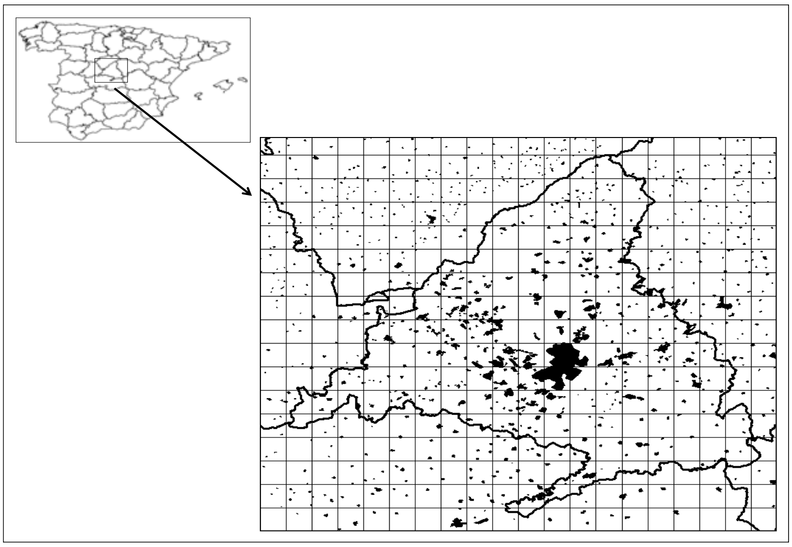


Figure 4.4: Parcels of land: Madrid

4.4.2 Actual *versus* Potential Output

In order to compare our predictions with the real data, we calculate the pattern of specialization by province, namely, how important is the production of a specific crop group over total production, considering the real scenario (Figure 4.5) and the potential scenario (Figure 4.6).

First we focus on the actual pattern of specialization (Figure 4.5). The provinces located in the center (namely, Burgos, Cuenca, Guadalajara, Madrid, Navarra, Palencia, Segovia, Soria, and Teruel) specialized in cereals (panel 1), while the provinces located in the northwest (A Coruña, Asturias, Cantabria, Lugo, Pontevedra, and Ourense) specialized mainly in tubers (panel 3). Ávila, Cádiz, and Valladolid specialized in the production of industrial crops (panel 4). The provinces that specialized in fodders (panel 5) were located in the north (Cantabria, Guipúzcua, and Vizcaya) and the northeast (Huesca, Girona, Lleida, and Zaragoza). Almería, Las Palmas de Gran Canarias, and Santa Cruz de Tenerife specialized in the production of vegetables (panel 6). The Mediterranean regions (mainly, Alicante, Castellón, and Valencia) specialized in the production of citrus (panel 7), Ciudad Real in the production of grapes (panel 8), and Jaén in the production of olives (panel 9).[34]

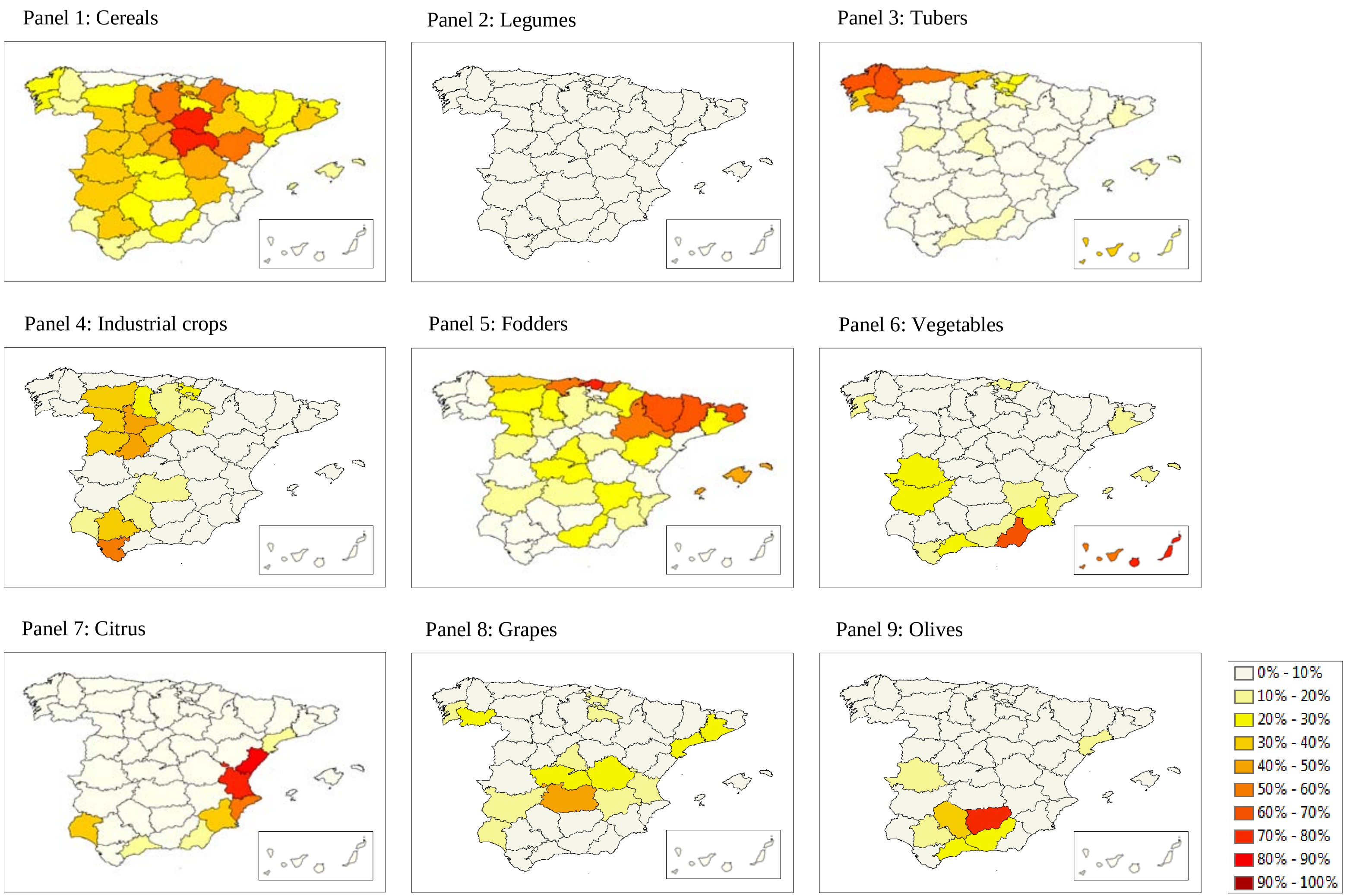
Nevertheless, according to our predictions presented in Figure 4.6, we observe that most

[34] Appendix 4.11 reports these figures by subperiod in order to check whether the actual pattern of specialization by province changed over time. Nevertheless, if we aggregate the data across provinces, the pattern of specialization in Spain during the period 1975–2011 was the following: cereals (29.41 percent), fodders (20.21 percent), industrial crops (13.31 percent), vegetables (8.24 percent), grapes (8.06 percent), citrus (7.34 percent), tubers (6.72 percent), olives (5.95 percent), and legumes (0.25 percent). (Author's computation using the AEA information.)

of the provinces, the ones situated in the center, south, and the Mediterranean and the two provinces that form the Canary Islands, should have specialized in the production of olives (panel 9). Asturias and Ourense should have specialized in the production of cereals (panel 1), and the northeast provinces, Barcelona, Girona, and Teruel, in legumes (panel 2). Finally, a group of provinces (Baleares, Cádiz, Ciudad Real, Huelva, Lleida, and La Rioja) should have specialized in the production of grapes.

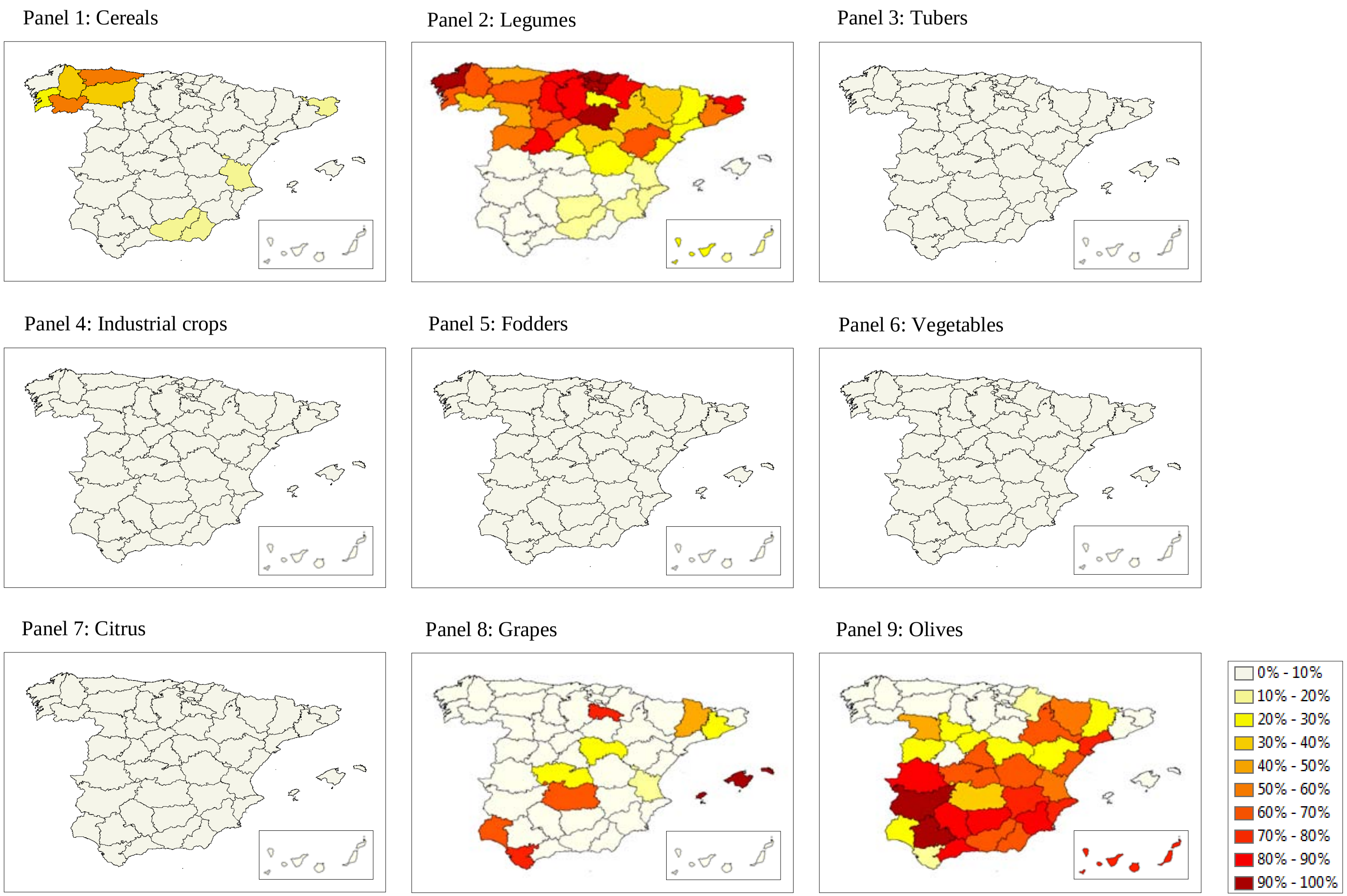
Therefore, our potential scenario predicts that most provinces should have specialized in legumes and olives. This prediction could be due to the price effect, namely, the real prices of legumes and olives were considerably higher than the prices of the other crops (see Figure 4.3). Under these circumstances, the profitability (namely, the crop revenues) obtained with these products was higher than with the ones that were expected to be most suitable.

Figure 4.5: Actual pattern of specialization by province, 1975–2011



Source: AEA (MAGRAMA). Calculus: Author.

Figure 4.6: Potential pattern of specialization by province, 1975–2011



Source: Own elaboration using data from the GAEZ project.

4.4.3 Econometric Specification

In order to quantify the relationship between actual and potential output, we propose an econometric exercise. The key issue is understanding whether the measures implemented by the CAP affected the relationship between these two variables.

Following Costinot and Donaldson (2012), we regress the logarithm of actual output on the logarithm of potential output.[35] Then the general econometric specification is defined as follows:

$$\ln(\text{actual output})_{rct} = \alpha + \beta \ln(\text{potential output})_{rct} + \gamma_r + \delta_c + \theta_t + u_{rct} \ , \tag{4.5}$$

where α is the constant, β is the elasticity between actual and potential output, γ_r is province fixed effects, δ_c is crop fixed effects, θ_t is time fixed effects, and u_{rct} is the error term. Our coefficient of interest is β. A value close to one means that the current production system converges near the potential one, indicating that real agricultural production is close to the most efficient one.

The previous estimation delivers an average slope coefficient (β) that does not allow us to assess whether important changes occurred over time. Therefore, we propose another technique in which a regression is run every year: there is a year-by-year comparison between actual and potential output.[36] Then, for year t, we define the following econometric specification:

$$\ln(\text{actual output})_{rc} = \alpha + \beta_t \ln(\text{potential output})_{rc} + \gamma_r + \delta_c + u_{rc} \ , \tag{4.6}$$

where $t = 1975, ..., 2011$. To ease its interpretation, the estimated coefficient (β_t) and its corresponding 5 percent confidence interval will be presented graphically. Then, we are able to follow the trend of β_t over time and to detect whether it has been affected by the adoption of the CAP rules and its corresponding reforms.

However, in order to better capture the effects of the policy, we need to separate the crops that were regulated since the beginning of the CAP from the others. To this extent, we create a dummy variable called *r_crops* to distinguish these two groups of crops.[37] Then we include this dummy and its interaction with potential production (in logarithms) in the econometric

[35] Logarithms serve to control in part for heterogeneity. In order to avoid the missing values observations, we add one unit to all observations before taking logarithms. Costinot and Donaldson (2012) use this technique as a robustness check.

[36] Namely, we run several cross-section regressions instead of exploiting the panel data structure.

[37] The dummy variable *r_crops* includes 11 crops: barley, maize, oat, rice, rye, sorghum, wheat, sugar beet, grape to transformation, olives, and olive oil.

specification:

$$
\begin{aligned}
\ln(\text{actual output})_{rct} &= \alpha + \beta \ln(\text{potential output})_{rct} + \beta^{r\text{-}crops} \ln(\text{potential output})_{rct} \cdot r\_crops + \\
&\quad + r\_crops + \gamma_r + \delta_c + \theta_t + u_{rct} \ , \qquad (4.7)
\end{aligned}
$$

where $\beta^{r\text{-}crops}$ is the elasticity between actual and potential output of those crops that were historically more regulated. A positive coefficient means that the crops that were regulated initially are closer to their efficient scenario with respect to the later regulated crops (which are in the control group).

One weakness of specification (4.7) could be related to the fact that the crops that were regulated since the beginning of the CAP were not selected exogenously. Then, the dummy *r_crops* could be correlated with some unobservable variables (included in the error term) that also affect actual output. If so, one should expect that initially regulated crops recorded lower levels of productivity at the beginning of the 1960s and higher productivity growth over time.

Using data from the AEA, Figure 4.7 depicts actual productivity in four different panels, one for each subperiod. The x-axis enumerates the 25 crops, and the y-axis represents the median (actual) productivity measured in tons per hectare (t/ha).[38] This figure allows us to analyze whether initially regulated crops were the least productive ones and whether their productivity followed a different trend with respect to the later regulated crops.

We observe that in the 1970s, a group of crops records very low productivity levels (less than 10 t/ha). Within this group there are crops that were regulated by the CAP since the beginning (barley, maize, oat, rice, rye, sorghum, wheat, olives, and olive oil) and crops that were not (dry bean, chickpea, dry pea, sunflower, and tobacco). Nevertheless, the sugar beet, an initially regulated crop, is one of the most productive crops (30 t/ha). Therefore, in our sample we cannot take for granted that initially regulated crops were the least productive at the beginning of the period.

Concerning the productivity trend over time, among the crops that were regulated by the CAP since the beginning, we observe a slight improvement in the productivity of maize and a considerable improvement in the productivity of sugar beet (from 30 t/ha to 70 t/ha). As for other less regulated crops, such as vegetables (cabbage, carrot, onion, and tomato), we also detect an improvement in their real productivity. Thus, we can also conclude that not only

[38] Median productivity refers to the representative province.

initially regulated crops improved their productivity in time.

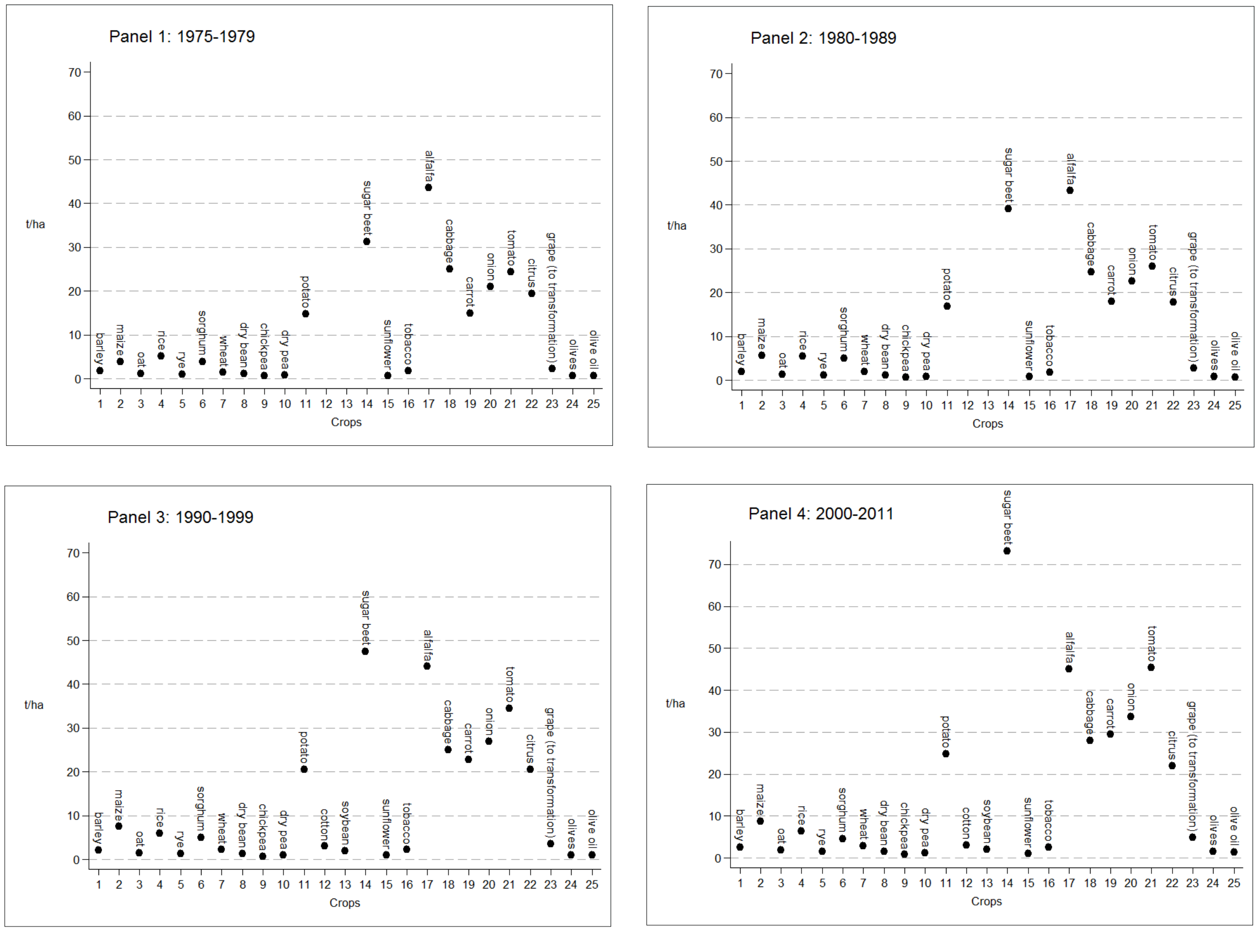
Overall, the previous results show that in Spain, the real productivity of the crops that were regulated since the beginning of the CAP was not different from the real productivity of the other crops. Moreover, according to Fearne (1997), the design of the CAP was strongly influenced by the national agricultural policies of the EEC-6. Since Spain joined the EEC in 1986, one could reasonably assume that the dummy *r_crops* is exogenous for Spain.

Finally, in order to detect the effects of the CAP during the period of study, again we propose an exercise in which a regression is run every year t:

$$\begin{aligned} \ln(\text{actual output})_{rc} \quad = \quad & \alpha + \beta_t \ln(\text{potential output})_{rc} + \beta_t^{r\_crops} \ln(\text{potential output})_{rc} \cdot r\_crops + \\ & + r\_crops + \gamma_r + \delta_c + u_{rc} \ . \end{aligned} \tag{4.8}$$

where $t = 1975, ..., 2011$. For the sake of simplicity, the estimated coefficient $(\beta_t^{r\_crops})$ and its corresponding 5 percent confidence interval will be presented graphically.

Figure 4.7: Actual productivity by crop and subperiod



Source: AEA (MAGRAMA, 2013). Calculus: Author.

Notes: The y-axis corresponds to the median (actual) productivity measured in tons per hectare (t/ha). Crops are ordered and numbered taking into account the crop groups presented in Table 1: cereals (1-7), legumes (8-10), tubers (11), industrial crops (12-16), fodders (17), vegetables (18-21), citrus (22), grapes (23), and olives (24 and 25). Information about cotton (12) and soybean (13) is not available until 1992.

4.5 Results

In order to better interpret the econometric results, first we map the real production patterns in the Spanish agricultural sector across time.

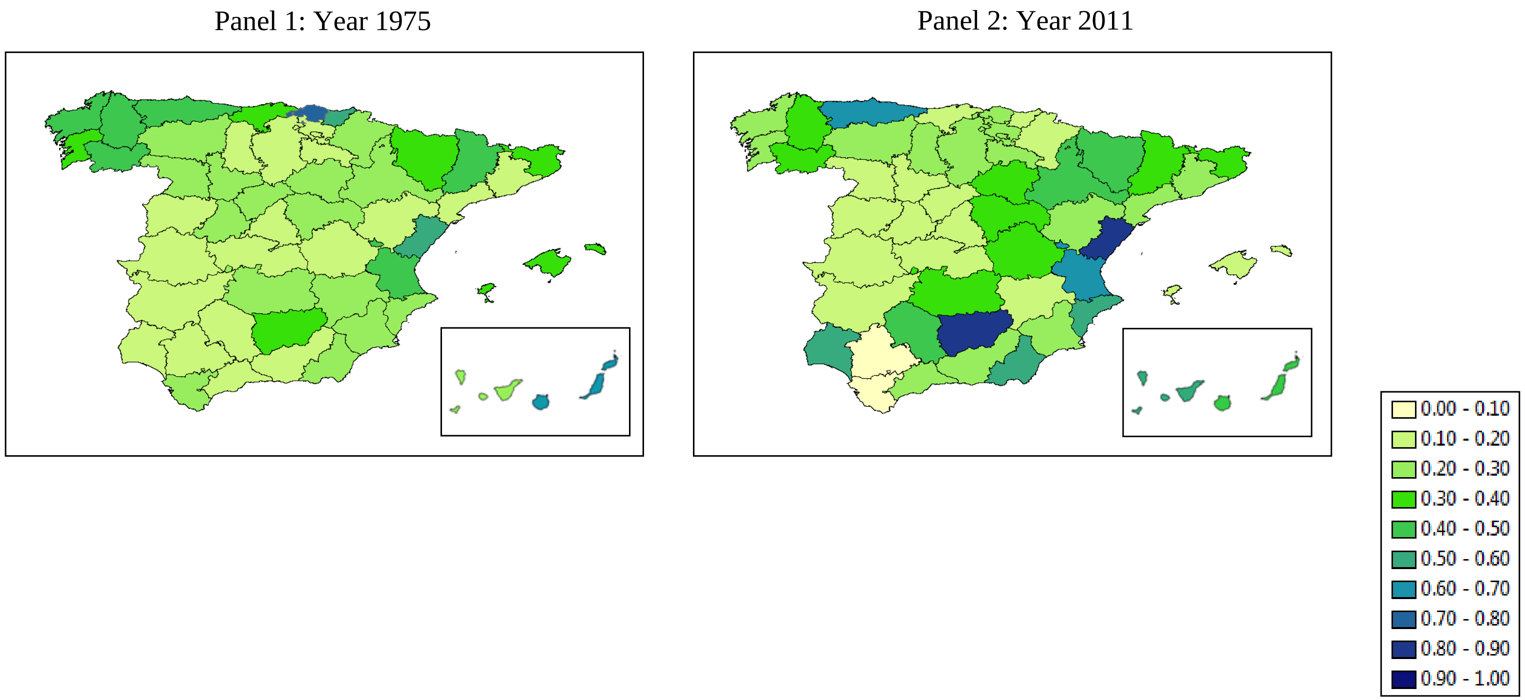
We achieve this objective by computing the Herfindahl-Hirschman index (henceforth HH index). This index provides information about the degree of specialization in production, and it is computed as the sum of the square of the shares of each crop in total agricultural output (namely, $HH = \sum_{c=1}^{25}(\frac{q^c}{Q})^2$, where q^c is the output of crop c and $Q = \sum_{c=1}^{25} q^c$ is total output). The index takes values between zero and one. Values close to zero indicate a low degree of specialization (in other words, diversification in production), while values close to one mean a high degree of specialization.

Figure 4.8 represents the HH index at the province level in 1975 and 2011. Looking at this picture, we can easily detect whether Spanish agriculture became more specialized as a consequence of the progressive European integration process and associated policies. According to Donald et al. (2002), European farmers have been encouraged to increase productivity by becoming more specialized, resulting in a polarization of production areas and a loss of mixed farming.

We observe that the provinces situated in the east—the Mediterranean area—achieve higher HH values in 2011 than in 1975, indicating a larger degree of specialization at the end of the period. Specifically, the index increases in 24 provinces, is relatively constant in 8 provinces, and decreases in 16 provinces. (See Table 4.8 in Appendix 4.11 for quantitative details.)

Overall, we observe an increase in the specialization in production by province over time, but we need to rely on the econometric exercise because we are interesting in assessing the existence of potential effects associated with the adoption of the CAP.

Figure 4.8: Herfindahl-Hirschman index by province, 1975 and 2011



Source: AEA (MAGRAMA). Calculus: Author.

Following the general econometric specification (4.5), we regress the logarithm of actual output over the logarithm of potential output. Our coefficient of interest is a measure of the correlation between these two variables. A high correlation means that actual output is close to potential output, whereas a low correlation means that actual output is far from potential output.[39]

Econometric outputs are presented in Table 4.2.[40] Column 1 refers to the baseline specification, without controlling for unobservable characteristics, while columns 2 and 3 include fixed effects. Column 2 includes individual fixed effects, namely, province, crop, and time dummy variables. Province dummy variables control for specific spatial, time-invariant characteristics, such as the weather or type of soil. Crop dummy variables control for specific, time-invariant characteristics related to each crop (for instance, the GAEZ agronomic models could be more accurate for some crops than for others). Time dummy variables control for economic shocks that could affect all crops and provinces. Finally, column 3 presents heavily constrained estimations by including bilateral fixed effects, namely, province-time and crop-time dummy variables. The former control for natural phenomena that affect the productivity of all the crops in a specific province in some years (for instance, freezes during the winter or fires during the summer), while the latter control for crop-specific regulations and policies that are time-variant. The inclusion of high-dimensional fixed effects could be crucial to solving the omitted variables problem.[41]

According to our findings, the estimated coefficient ($\widehat{\beta}$) is positive and statistically significant in the three specifications. Its magnitude is equal to 0.08 in columns 1 and 3, and 0.06 in column 2. Therefore, our first set of results identifies a positive (though weak) relationship between the actual and potential output.[42]

In order to track the evolution of the level of efficiency achieved by the Spanish agricultural sector, we estimate specification (4.6) by year. Then, instead of obtaining an average estimated coefficient ($\widehat{\beta}$), we get an estimated time-specific coefficient ($\widehat{\beta}_t$).

The estimated coefficient ($\widehat{\beta}_t$) and its corresponding 5 percent confidence interval are presented in Figure 4.9. We observe that from 1986 to 1992, the estimated coefficient follows a

[39] *A priori* we do not know whether actual output is above or below its benchmark, then the inefficiency could be due to overproduction (actual output is above its benchmark) or underproduction (actual output is below its benchmark).

[40] All regressions include standard errors that are clustered by province to account for potential within-province (across crops) interdependence.

[41] To run this regression, we employ the program *reg2hdfe* in Stata (StataCorp, 2012), implementing the algorithm developed by Guimarães and Portugal (2010). It allows us to control for high-dimensional fixed effects without incurring storage problems. But the value of the constant is not included in the statistical output.

[42] We compare our results with the ones obtained in Costinot and Donaldson (2012), taking into account that the spatial dimension of our exercises is very different. Their preferred coefficient is equal to 0.212, though it decreases to 0.096 after including country dummy variables.

Table 4.2: Actual *versus* potential outputs

Dependent variable:	ln(actual output)		
	(1)	(2)	(3)
ln(potential output)	0.0834***	0.0656***	0.0814***
	(0.0215)	(0.0180)	(0.0217)
constant	11.9720***	4.1079***	
	(0.3711)	(0.6934)	
Fixed effects:			
province		X	
crop		X	
time		X	
province-time			X
crop-time			X
Observations	42,350	42,350	42,350
R-square	0.0080	0.5421	0.5634

Estimator: Ordinary least squares.

Standard errors clustered by province are in parentheses.

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

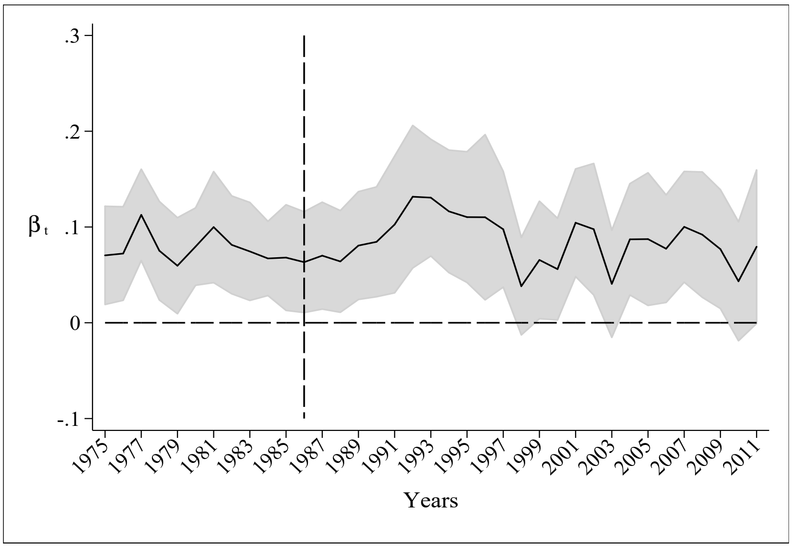
positive trend from $\widehat{\beta}_{1986} = 0.06$ to $\widehat{\beta}_{1992} = 0.13$ (and the difference between these two coefficients is statistically significant). This means that the difference between actual and potential output decreases over time. Therefore, we can confirm that the entry to the EEC and the adoption of the CAP policies boosted efficiency in the Spanish agricultural sector.

Unfortunately, we cannot distinguish which part of this efficiency improvement is explained by being part of a common market and which part by the introduction of the CAP measures. One of the benefits of being part of a common market is the increase in trading among partners (namely, trade creation effects). This means that each country specializes in the goods that it produces more efficiently and imports the remaining ones. Indirectly, a country could experience an improvement in the allocation of resources, which in our case is reflected through land use. Alternatively, the implementation of specific CAP measures could have affected farmers' production incentives.

Although from 1992 to 1996 the estimated coefficient follows a decreasing trend, the estimated values are statistically higher than the ones obtained before 1986. Instead, from 1997 onward, the estimated values are not statistically different from the ones obtained before 1986.

Therefore, these results suggest that the entry of Spain in the EEC and the implementation of the CAP measures had some effects that vanished over time, while the 2003 CAP reform did not have an impact on the production decisions. At this point, it could be interesting to discriminate among crops to assess the true effectiveness of the policy.

Figure 4.9: Actual *versus* predicted output by years



Following specification (4.7), we include some new variables in our general econometric specification in order to check whether there are important differences, in terms of efficiency, between the crops that were historically more regulated by the policy and the others. Estimation results are in Table 4.3. As before, we run three specifications that differ according to the inclusion of a set of dummy variables that control for unobservable characteristics. We are interested in the estimated coefficient $\widehat{\beta}^{r\_crops}$, which reflects the relationship between the actual and potential output of the crops that were historically more regulated. The estimated coefficient is positive and statistically significant in the three columns (0.17, 0.15, and 0.17, respectively), meaning that initially regulated crops are closer to their benchmark if compared with the later regulated crops (control group).[43]

Again we exploit the temporal dimension of our sample by running a year-by-year regression according to the econometric specification (4.8). We represent the estimated coefficient ($\widehat{\beta}_t^{r\_crops}$) and its 5 percent confidence interval over time (Figure 4.10). Again, we observe that the entrance of Spain to the EEC and the implementation of the CAP policies affected efficiency, since the estimated coefficient follows a jump from $\widehat{\beta}_{1986}^{r\_crops} = 0.0811$ to $\widehat{\beta}_{1987}^{r\_crops} = 0.1820$, and the

[43]It does not mean that the actual production of the regulated crops is higher than the actual production of the nonregulated crops. Remember that the coefficient only provides information about how far actual output is from its benchmark, potential output.

Table 4.3: Actual *versus* potential outputs: regulated crops

Dependent variable:	ln(actual output)		
	(1)	(2)	(3)
ln(potential output)	-0.0018	-0.0130	-0.0204
	(0.0244)	(0.0210)	(0.0316)
ln(predicted output)·r_crops	0.1659***	0.1539***	0.1731***
	(0.0300)	(0.0333)	(0.0416)
r_crops	0.1288	-2.9031*	omitted
	(0.3330)	(1.5312)	
constant	11.9720***	4.1079***	
	(0.3711)	(0.6934)	
Fixed effects:			
province		X	
crop		X	
time		X	
province-time			X
crop-time			X
Observations	42,350	42,350	42,350
R-square	0.0186	0.5452	0.5665

Estimator: Ordinary least squares.
Standard errors clustered by province are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
r_crops: barley, maize, oat, rice, rye, sorghum, wheat, sugar beet, grape to transformation, olives, and olive oil.

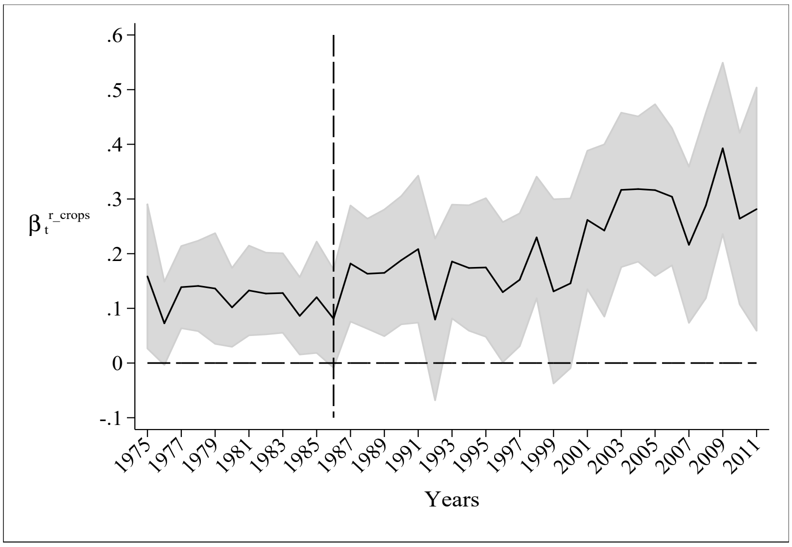
difference between these coefficients is statistically significant.

Furthermore, in Figure 4.10 we obtain a new and interesting result. Since 2000 we observe that the estimated coefficient follows a positive trend, which means that initial regulated crops are closer to their efficient scenario with respect to later regulated crops (the control group).[44] In addition, these estimated values are significantly higher than the ones obtained before 1986. Then, we detect an improvement in the efficiency level for a sample of specific crops. This improvement could be an effect of the Fischler reform (2003), where direct payments were decoupled from the level of output. Under this perspective, the incentives for increasing production of the regulated crops in order to enjoy the benefits of the communitarian policy disappeared, and consequently actual production of these crops approached potential production.

[44] The peak is reached in 2009, where $\widehat{\beta}_{2009}^{r\_crops} = 0.39$.

Therefore, through specification (4.8), we provide evidence that the communitarian policy had a tangible impact on the production decisions in Spain. The 2003 CAP changes, mainly a more significant decoupling of payments from production, produced greater market orientation and lower market distortions (OECD, 2004).

Figure 4.10: Actual *versus* predicted output by years: regulated crops



4.6 Robustness Checks

We find that the entry of Spain to the EEC and the Fischler reform (2003) had an impact on the efficiency level achieved by Spanish agricultural production. Nevertheless, these results rely on the definition of an efficient scenario, which is used as a benchmark. In order to overcome this critique, we run some robustness checks by varying some of the key variables that were used to compute the benchmark.

Our strategy is to deal with three new scenarios. In the first robustness exercise, we select another GAEZ variable as a proxy for potential productivity. In the second robustness exercise, we replace Spanish market prices with those of other countries. Finally, in the third robustness exercise, we redefine the land area devoted to agricultural production.

4.6.1 Potential Productivity

The first robustness exercise consists of using a new variable as a proxy for potential productivity (A_f^c). This is a key variable to determine both efficient land allocation and potential output.

We replace the PPC variable with the *total production capacity* (TPC) variable.[45]

We repeat the general econometric estimation (4.5) and results are presented in Table 4.4.[46] The estimated coefficient ($\widehat{\beta}$) is positive and statistically significant in the three columns, which means that our results are consistent to the change of one of the key variables.

Table 4.4: Actual *versus* potential output: potential productivity

Dependent variable:	ln(actual output)		
	(1)	(2)	(3)
ln(potential output)	0.1144***	0.0486***	0.0649***
	(0.0202)	(0.0169)	(0.0199)
constant	11.8360***	5.5296***	
	(0.3713)	(0.3895)	
Fixed effects:			
province		X	
crop		X	
time		X	
province-time			X
crop-time			X
Observations	43,150	43,150	43,150
R-square	0.0128	0.5409	0.5634

Estimator: Ordinary least squares.

Standard errors clustered by province are in parentheses.

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

[45]The TPC reports potential production capacity in terms of output density (kg/ha) by grid-cell. This is the variable used in the Costinot and Donaldson (2012) empirical exercise. We do not use this variable in the main analysis because it prevents us from taking into account the adoption of the irrigation technique that evidence identifies as important in Spain.

[46]The number of observations is higher because the TPC variable has information for rape. According to the classification presented in Table 4.1, this crop belongs to the group industrial crops.

4.6.2 Market Prices

The second robustness exercise considers alternative market prices in order to assess whether our results are driven by nominal values.[47] We repeat the analysis by replacing the Spanish agricultural prices with an average between the Greek and Portuguese agricultural prices. We select these countries because they joined the EEC during the 1980s and their weather and soil characteristics are similar to those in Spain.

In order to perform this exercise, we extract data from the FAOSTAT database *Producer prices*.[48] Unfortunately, this database does not contain information about all the crops used in our analysis. Therefore, we reduce our sample from 25 to only 19 crops.[49]

Results are presented in Table 4.5. We observe that the estimated coefficient ($\widehat{\beta}$) is positive and significant in the three columns, though its magnitude is smaller than the one obtained in the main analysis, particularly in the specifications that include fixed effects. This decrease is expected because we are computing the efficient scenario of the Spanish agricultural sector using external market prices, whereas the actual scenario captures the effects of actual Spanish market prices. But again, we can confirm that our main results are robust to the variation of an important variable.

4.6.3 Land Endowment

Finally, we focus on the whole size of land devoted to agriculture. Previously, we use all available land, except urban areas, to compute potential output. Here we consider the actual area according to a specific definition elaborated by an EU project.

In 1985 the EU launched the coordination of information on the environment (CORINE) program with the objective of compiling and coordinating information about the environment and natural resources in the European territory. One of the projects included in this program was CORINE Land Cover (CLC), which was taken over by the European Environmental Agency (EEA) in 1995. The objective of this project was to build a European database on effective land cover uses.[50]

At the moment we have information for only three years: 1990, 2000, and 2006. In order to

[47] The role of the prices in the numerical exercise is very important. A drastic change in relative prices could affect the land allocation across crops, and indirectly this would affect potential output.

[48] To have all prices in the same nominal currency (euro/ECU), we use bilateral exchange rates reported by Eurostat.

[49] We do not have information for the following crops: alfalfa, cotton, olives, dry pea, soybean, and sorghum.

[50] Land cover nomenclature comprises three levels according to the level of precision: the first level considers 5 items, the second level considers 15 items, and the third level considers 44 items (for additional information, see EC, 1995, p. 21). In this exercise we use the first level, specifically the item called "Agricultural areas."

Table 4.5: Actual *versus* potential output: market prices

Dependent variable:	ln(actual output)		
	(1)	(2)	(3)
ln(predicted output)	0.0761***	0.0151*	0.0239*
	(0.0180)	(0.0088)	(0.0128)
constant	12.8293***	3.7241***	
	(0.3671)	(0.5618)	
Fixed effects:			
province		X	
crop		X	
time		X	
province-time			X
crop-time			X
Observations	33,250	33,250	33,250
R-square	0.0097	0.4703	0.4939

Estimator: Ordinary least squares.

Standard errors clustered by province are in parentheses.

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

compute the new benchmark, we need to extend the available information on effective land use throughout our whole period. As such, we are forced to introduce an operative hypothesis. The 1990 database (namely, the land area devoted to agriculture in 1990) is exploited to calculate potential production for the period 1975–1999, the 2000 database for the period 2000–2005, and the 2006 database for the period 2006–2011.

Results are presented in Table 4.6. Once more, the estimated coefficient ($\widehat{\beta}$) is positive and significant in the three columns, which means that our results are also robust to a change in selected land areas.

Table 4.6: Actual *versus* potential output: land endowment

Dependent variable:	ln(actual output)		
	(1)	(2)	(3)
ln(predicted output)	0.0824***	0.0719***	0.0901***
	(0.0232)	(0.0188)	(0.0227)
constant	11.9947***	4.2089***	
	(0.3689)	(0.6940)	
Fixed effects:			
province		X	
crop		X	
time		X	
province-time			X
crop-time			X
Observations	42,350	42,350	42,350
R-square	0.0070	0.5422	0.5635

Estimator: Ordinary least squares.

Standard errors clustered by province are in parentheses.

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

4.7 Conclusions

In this paper, we analyze the extent to which the CAP affected the level of agricultural production in Spain. To achieve our objective, we create a novel database by considering 50 provinces and 25 crops, and we compare the differences between current and potential output over the period 1975—2011.

Our results assess the improvement in production efficiency that occurred after Spain entered the EEC (1986) and especially after the Fischler reform and the decision to decouple money transfers from the level of output (2003).

One of the limitations of our empirical strategy is that although we observe an efficiency improvement in the agricultural sector during the second half of the 1980s, we are not able to disentangle the causes. It could be because Spain joined the common market—namely, the trade creation effects—or because the implementation of the CAP measures produced specific production incentives. Further investigation in this direction is recommended in order to capture the role of these two different effects and provide better insights with respect to policy conclusions.

However, this contribution provides evidence that the 2003 CAP reform, which broke the link between subsidies and production, had a positive impact on real production in Spain. This impact could be due to a change in farmers' production incentives, since they became more efficient as they relied more on market characteristics rather than just on subsidy requirements. In light of these results, any future CAP modification or reform should be designed by taking into account this policy scheme to maintain the efficiency target: lump-sum money transfers independent from the output level or other pecuniary incentives.

As for our research strategy, we propose two further extensions, the former to overcome the limitation previously discussed and the latter to obtain a much more complete evaluation of the policy.

In order to distinguish the impact of the CAP measures from the effects of being part of a common market, we can extend the analysis to other industrial sectors (which were also affected by the entrance of Spain to the EEC). The idea is to estimate the potential effects of being part of a common market, in terms of production level and trade flows, for each sector separately as well as for the whole economy (an average effect). Then, we can compare this average effect with the results obtained by each sector. This comparison is especially interesting in the case of agriculture, since part of the differences (if any) could be explained by the implementation of the CAP. Therefore, through this exercise we could identify the effective impact of the CAP measures in Spanish agricultural production, particularly during the second half of the 1980s.

A further extension would be to enlarge the sample of analysis by including other countries both inside and outside the EU. The rationale is to gather information by referring to other case studies in order to assess whether they also experienced an improvement in the level of production and, if so, compare the magnitude of these improvements with the Spanish one. Under this perspective, a second interesting question to address is checking whether the EU membership (and therefore the possibility of enjoying the CAP policy) has truly been a discriminating device to record a larger improvement in the performance of agricultural production. This extension, which is very challenging in terms of data compilation, will allow us to elaborate a deeper evaluation of the CAP, since it will include a larger spatial dimension.

4.8 Appendix A

Our database is built by considering two main data sources: potential productivity values are taken from the GAEZ project elaborated by the FAO and IIASA, while prices and actual outputs are taken from the AEA published by the MAGRAMA. Whenever data (concerning potential productivity and price) are not available, we need to approximate them. This appendix outlines our strategy for managing the information from these two sources in order to obtain an adequate database for implementing the numerical exercise described in Appendix 4.10.

4.8.1 Representative Value: Citrus

We apply a special adjustment technique for citrus. The GAEZ reports information about *citrus,* while the AEA reports more disaggregated information (the group of citrus is composed of oranges, mandarins, and lemons). Using the AEA information, we calculate a representative value for citrus.

First, using information about quantities (q) and prices (p), we calculate the total revenue for citrus in province r and year t:

$$Revenue_{rt}^{citrus} = q_{rt}^{o} p_t^{o} + q_{rt}^{m} p_t^{m} + q_{rt}^{l} p_t^{l} \ ,$$

where o refers to oranges, m refers to mandarins, and l refers to lemons.

Then, for each year t we define the representative price of citrus as

$$p_t^{citrus} = p_t^{o} \left(\frac{q_t^{o}}{q_t^{o} + q_t^{m} + q_t^{l}} \right) + p_t^{m} \left(\frac{q_t^{m}}{q_t^{o} + q_t^{m} + q_t^{l}} \right) + p_t^{l} \left(\frac{q_t^{l}}{q_t^{o} + q_t^{m} + q_t^{l}} \right) ,$$

where p_t^i $(i = o, m, l)$ is the price in year t and $q_t^i = \sum_{r=1}^{50} q_{rt}^i$ $(i = o, m, l)$ is the national production in year t. The relative weight of the price of each product in the representative price of citrus p_t^{citrus} is measured according to the weight that each variety has in total production.

Finally, using total revenue and representative price, we calculate a representative output of citrus by province r and year t:

$$q_{rt}^{citrus} = \frac{Revenue_{rt}^{citrus}}{p_t^{citrus}} .$$

This output is considered to be the real one and will be compared with the predicted one.

4.8.2 Prices

Prices are needed in order to determine the efficient land allocation. Unfortunately, prices are not available for some specific crops and years. We implement different strategies to approximate the missing values in the AEA.

Since the AEA and FAOSTAT databases do not have many differences, the first method consists of complementing the missing values with data reported by the FAOSTAT database *Producer prices*. We use information from the FAOSTAT for the following crops and years: oranges (1976–1979), lemons (1976–1979), table olives (1976–1979), potatoes (1986), and carrots (1975–1984).

However, when the information is not available in FAOSTAT, we use a second method: we calculate an average value of prices, taking into account the last and the first available prices of the period with missing information. This method is implemented for alfalfa and mandarins. The average price is calculated using the prices reported for the years 1975 and 1980. Then, this average is used to cover the period 1976–1979.

The third method consists of determining a representative group. We implement it by comparing the prices of two crops that belong to the same category (for instance, table olives and olive oil), where one variety possesses complete information and the other has none. The strategy is to use the available prices to calculate a ratio between the prices of the two varieties. Then, this ratio is exploited as a proportional factor to complete the missing values in the series.

Crops to which we apply this adjustment are olive oil and grape to transformation. The price of olive oil is available since 1985. In order to complete the period (1975–1984), we use the prices of table olives as a reference. We focus on the available prices during the 1980s (that is, from 1986 to 1989). For each year, a ratio between the price of table olives and the price of olive oil is established. The ratio in year t is defined as $r_t^{p(oo)} = \frac{p_t^{olive}}{p_t^{olive\_oil}}$. We get the following ratios: $r_{85}^{p(oo)} = 0.58$, $r_{86}^{p(oo)} = 0.26$, $r_{87}^{p(oo)} = 0.22$, $r_{88}^{p(oo)} = 0.32$, and $r_{89}^{p(oo)} = 0.21$. An average of these ratios is calculated and is equal to 0.32. Then, the price of olive oil in year t is equal to $p_t^{olive\_oil} = \frac{p_t^{olive}}{0.32}$ for the period 1975–1984 (namely, the price of table olives in year t over the average ratio).

For grape to transformation, we perform a similar exercise but, in this case, we use the prices of table grapes as a reference group. The price of grape to transformation is available by varieties and regions but only for three years: 1995, 2000, and 2008.[51] We choose a representative variety

[51] Information about prices are extracted from Lissarrague García-Gutiérrez and Martínez de Toda Fernández (2009, p. 70, table 31). According to the previous study, the price source is *La Semana Vitivinícola*, N^o 2565 (1995), 2826 (2000), 3240 (2008).

(*tempranillo*) and calculate an average price, one per year.[52] Then we compare the prices of table grapes and grape to transformation and calculate a ratio for each year. The ratio in year t' is defined as $r_{t'}^{p(gg)} = \frac{p_{t'}^{grape}}{p_{t'}^{grape\_transf}}$, $t' = 95, 00, 08$. We get the following ratios: $r_{95}^{p(gg)} = 0.74$, $r_{00}^{p(gg)} = 0.65$, and $r_{08}^{p(gg)} = 1.13$. Using these ratios and the prices of table grapes, we estimate a price for grape to transformation. The price of grape to transformation in year t is equal to $p_t^{grape\_transf} = \frac{p_t^{grape}}{r_{t'}^{p(gg)}}$ where $t' = 95, 00, 08$. The ratio $r_{95}^{p(gg)}$ is used for the period $t \in [1975, 1996]$, the ratio $r_{00}^{p(gg)}$ is used for the period $t \in [1997, 2000]$, and the ratio $r_{08}^{p(gg)}$ is used for the period $t \in [2001, 2011]$.

4.8.3 Potential Productivity: Grapes

Grapes is a group composed of table grapes and grape to transformation (mainly for the production of wine). This group is important in Spanish agriculture. According to the AEA data, it represents around 6 to 8 percent of total agricultural production. Unfortunately, the GAEZ database does not report information for either of these two varieties. But it is relevant to include at least the grape to transformation in our analysis.[53] Therefore, we have to approximate potential productivity for the grape to transformation.

In order to calculate potential productivity, we choose another type of crop with similar characteristics: olives.[54] First, using information from the AEA, we calculate the *aggregate* actual productivity (kg/ha) for olives ($prod^{olive}$) and for grape to transformation ($prod^{grape\_transf}$).[55] Second, a ratio between the actual productivity of olives with respect to grape to transformation is calculated in order to establish a relation between both crops, that is, $r(og) = \frac{prod^{olive}}{prod^{grape\_transf}}$. Finally, using this ratio and the information reported by the GAEZ about olives (A_f^{olive}), we establish potential productivity for grape to transformation, one for each grid-cell f, $A_f^{grape\_transf} = \frac{A_f^{olive}}{r(og)}$. In this way, we can include this relevant crop in our sample of analysis to determine the efficient land allocation problem.

[52] According to *WineAccess Search*, *tempranillo* is arguably the most famous of Spain's native grapes. Source: http://www.wineaccess.com/wine/grape/tempranillo.

[53] In 1975 table grapes represented 8.87 percent of the total production of grapes, while the grape to transformation represented 91.13 percent. In 2011 this difference was even higher: table grapes represented 4.19 percent of the total production of grapes, while the grape to transformation represented 95.8 percent. (Author's computation using the AEA information.)

[54] Olives and grapes follow similar production systems, since they are very common crops in the Mediterranean area because of weather characteristics, temperate winters, and hot and dry summers (Espasa, 1998).

[55] *Aggregate* means that we consider all provinces $(\sum_{r=1}^{50})$ and all years $(\sum_{t=1}^{37})$.

4.9 Appendix B

The objective of this appendix is to provide an argument for the independence of the efficient factor allocation from demand conditions.[56]

In the case of Leontief technology, the efficient input allocation is independent from demand conditions and total output. The Leontief model considers an economy with C commodities and M elementary activities. An elementary activity is a combination of inputs to produce a good, such that each activity involves a different technique. But there is one commodity, say, the Cth, which is not produced by any activity but instead represents a primary factor (in our case, land). The technology has constant returns to scale, and there is no joint production.[57]

In the Leontief model *without substitution* possibilities, if a good should be produced by using a fixed combination of inputs, then there is one specific elementary activity for each good. In the Leontief model *with substitution* possibilities, however, a good may be produced with several techniques (for instance, $a_1, ..., a_{M_1}$ is the list of M_1 elementary activities capable of producing good c_1). Although the substitution of techniques is possible, it will never occur because it can be shown that a single optimal technique is associated with each good. Because of the constant returns to scale technology, the optimal technique is independent from demand conditions and the total level of output. This is known as the *nonsubstitution theorem* (Samuelson, 1951). The theorem depends critically on the presence of only one factor of production. With more than one factor of production, the optimal choice of technique should depend on the relative prices of these factors (see Chakravarty, 2005, p. 578).

Furthermore, according to the *First Welfare Theorem*, if a production vector $\overrightarrow{y} \in Y$ is profit maximizing for some strictly positive price vector $\overrightarrow{p} >> 0$, then $\overrightarrow{y}$ is efficient (here, the land allocation).

[56] For a more complete discussion, see Mas-Colell et al. (1995), chapter 5.

[57] Joint production occurs when more than one good is created sumultaneusly from one specific elementary activity.

4.10 Appendix C

This appendix provides a complete description of the numerical analysis implemented to determine the efficient land allocation and potential output for a specific province r.

(a) First, we have to compute the land revenue (Euros/ha) using information about the potential productivity and market prices. For year t and parcel of land f, the land revenue for each crop c is expressed as follows:

crop	revenue (Euros/ha)
1	$A_f^{c_1} p_t^{c_1}$
2	$A_f^{c_2} p_t^{c_2}$
...	
25	$A_f^{c_{25}} p_t^{c_{25}}$

The crop that maximizes the land revenue is the one that is chosen to be produced in parcel f. It is important to check *uniqueness*, that is, only one crop maximizes the land revenue.

(b) Given a province r, the sequence of computation at part (a) is repeated for all the parcels of land. For instance, in the case of Valencia ($r = 46$), 212 parcels are identified as $f_{8984}, f_{8985}, ..., f_{9195}$ (see Table 4.7).

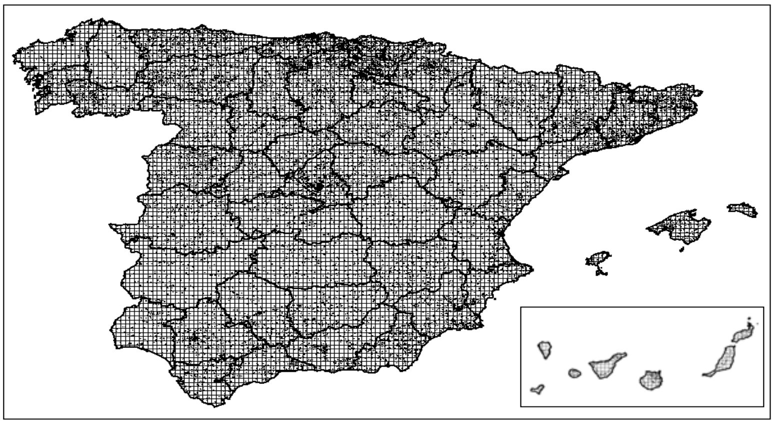
(c) The efficient land allocation is defined (Ω_{rt}^c): each crop c is produced in a set of parcels of land.

(d) The potential output of crop c in year t is calculated by taking into account the efficient land allocation, potential productivity, and the endowment, namely, $[Q_{rt}^c]^* = \sum_{f \in \Omega_{rt}^c} A_f^c \overline{L}_{rf}$.

The endowment $\overline{L}_{rf}$ is the total surface (ha) of each parcel of land f after excluding the urban areas (see Figure 4.11).

(e) Since prices are a time-variant variable, we need to repeat the actions of parts (a)–(d) for each year t ($t = 1975, ..., 2011$).

Figure 4.11: Parcels of land



Notes:
Spain is divided into 8,178 parcels.
Black areas represent urban areas (MAGRAMA).

Table 4.7: Provinces and grid-cells

Code	Province	Surface (ha)[1]	Surface (ha)[2]	id grid-cells[3]	Total grid-cells
1	Álava	303,511.75	296,298.39	[1, 84]	84
2	Albacete	1,491,726.15	1,477,252.79	[85, 362]	278
3	Alicante	581,978.64	571,776.96	[363, 483]	121
4	Almería	876,338.89	862,412.99	[484, 648]	165
5	Ávila	804,872.28	798,673.06	[649, 811]	163
6	Badajoz	2,179,020.85	2,167,784.96	[812, 1207]	396
7	Baleares	501,711.08	482,355.50	[1208, 1335]	128
8	Barcelona	775,678.54	743,692.05	[1336, 1491]	156
9	Burgos	1,428,085.43	1,409,511.47	[1492, 1786]	295
10	Cáceres	1,988,898.21	1,973,598.81	[1787, 2152]	366
11	Cádiz	744,090.26	733,281.59	[2153, 2295]	143
12	Castellón	663,644.66	657,513.80	[2296, 2431]	136
13	Ciudad Real	1,980,110.93	1,968,997.11	[2432, 2790]	359
14	Córdoba	1,376,898.90	1,361,251.81	[2791, 3042]	252
15	A Coruña	799,335.65	793,373.68	[3043, 3212]	170
16	Cuenca	1,712,891.89	1,702,891.36	[3213, 3524]	312
17	Girona	593,522.79	576,745.27	[3525, 3657]	133
18	Granada	1,263,509.25	1,252,104.11	[3658, 3887]	230
19	Guadalajara	1,220,292.53	1,208,379.84	[3888, 4125]	238
21	Guipúzcoa	198,093.95	190,603.62	[4126, 4176]	51
21	Huelva	1,014,602.14	1,007,739.09	[4177, 4366]	190
22	Huesca	1,564,748.98	1,549,344.46	[4367, 4667]	301
23	Jaén	1,348,633.00	1,337,381.50	[4668, 4911]	244
24	León	1,559,129.12	1,538,902.20	[4912, 5210]	299
25	Lleida	1,219,338.68	1,199,767.70	[5211, 5451]	241
26	La Rioja	504,152.31	496,182.65	[5452, 5563]	112
27	Lugo	988,042.80	983,008.71	[5564, 5758]	195
28	Madrid	802,532.89	758,441.09	[5759, 5927]	169
29	Málaga	730,686.24	718,275.30	[5928, 6076]	149
30	Murcia	1,131,311.86	1,115,461.11	[6077, 6288]	212
31	Navarra	1,038,561.22	1,023,830.61	[6289, 6505]	217
32	Ourense	729,351.27	721,670.09	[6506, 6658]	153
33	Asturias	1,060,902.89	1,042,849.76	[6659, 6871]	213
34	Palencia	804,946.92	797,018.60	[6872, 7040]	169
35	Las Palmas	419,167.63	408,387.45	[7041, 7135]	95
36	Pontevedra	451,187.03	445,047.56	[7136, 7242]	107
37	Salamanca	1,236,083.26	1,223,903.17	[7243, 7476]	234
38	Santa Cruz de Tenerife	351,564.52	342,252.51	[7477, 7553]	77
39	Cantabria	532,420.12	513,254.94	[7554, 7673]	120
40	Segovia	691,860.00	687,318.11	[7674, 7811]	138
41	Sevilla	1,404,454.96	1,387,309.96	[7812, 8069]	258
42	Soria	1,029,945.65	1,020,339.19	[8070, 8277]	208
43	Tarragona	631,536.56	624,117.81	[8278, 8409]	132
44	Teruel	1,481,064.10	1,472,030.85	[8410, 8691]	282
45	Toledo	1,536,034.62	1,522,083.87	[8692, 8983]	292
46	Valencia	1,080,997.96	1,066,068.36	[8984, 9195]	212
47	Valladolid	810,893.91	800,775.12	[9196, 9366]	171
48	Vizcaya	221,624.57	210,619.64	[9367, 9427]	61
49	Zamora	1,056,918.48	1,047,750.03	[9428, 9638]	211
50	Zaragoza	1,727,267.04	1,704,896.27	[9639, 9977]	339

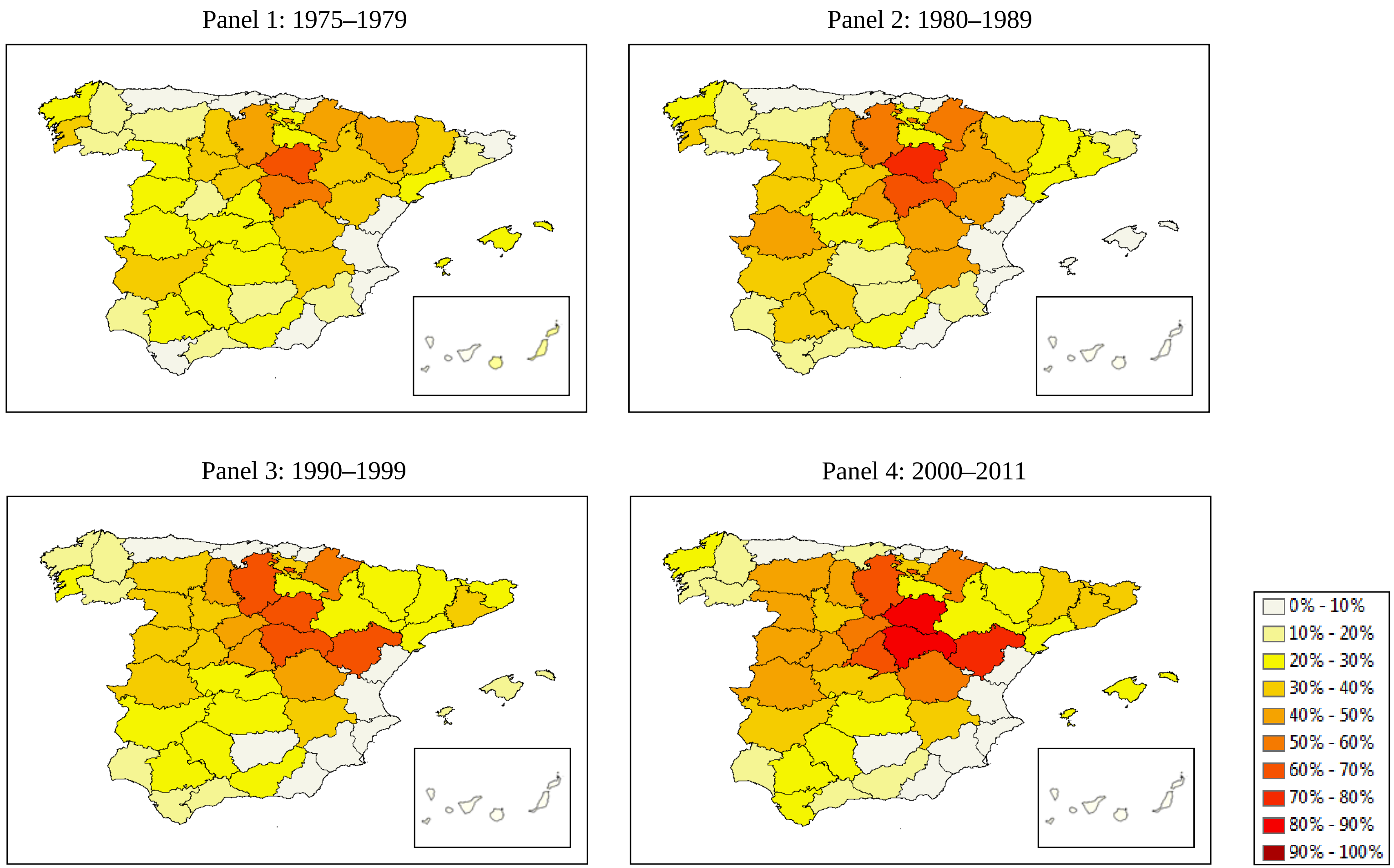
Notes:
[1, 2] Author's computation using the geographic software ArcGIS.
[2] Excluding urban areas (MAGRAMA, 2014).
[3] The identifier "id grid-cells" is defined by grid-cell and province. For this reason, there are 9,977 identifiers instead of 8,178 (total number of grid-cells in Spain).

4.11 Appendix D

We present the actual pattern of specialization by province and subperiod in Figures 4.12–4.20. We appreciate that there was an increase in the number of provinces specializing in the production of cereals, citrus, and grapes (namely, these crop groups represented more than 40 percent of total production). Nevertheless, there was a reduction in the number of provinces specializing in fodders.

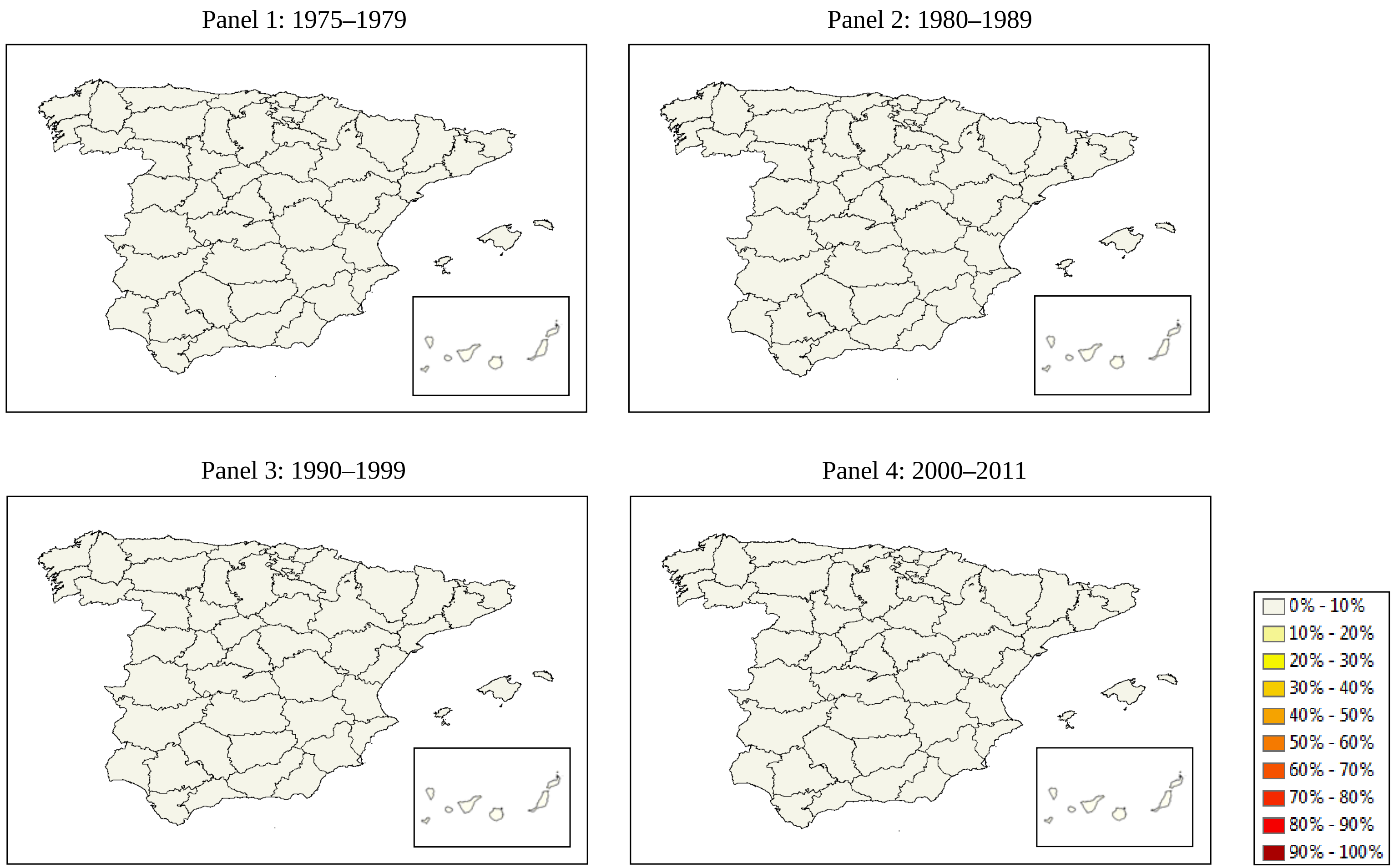
Table 4.8 reports the HH values at the province level in 1975 and 2011 in order to complement Figure 4.8.

Figure 4.12: Specialization pattern by province: cereals



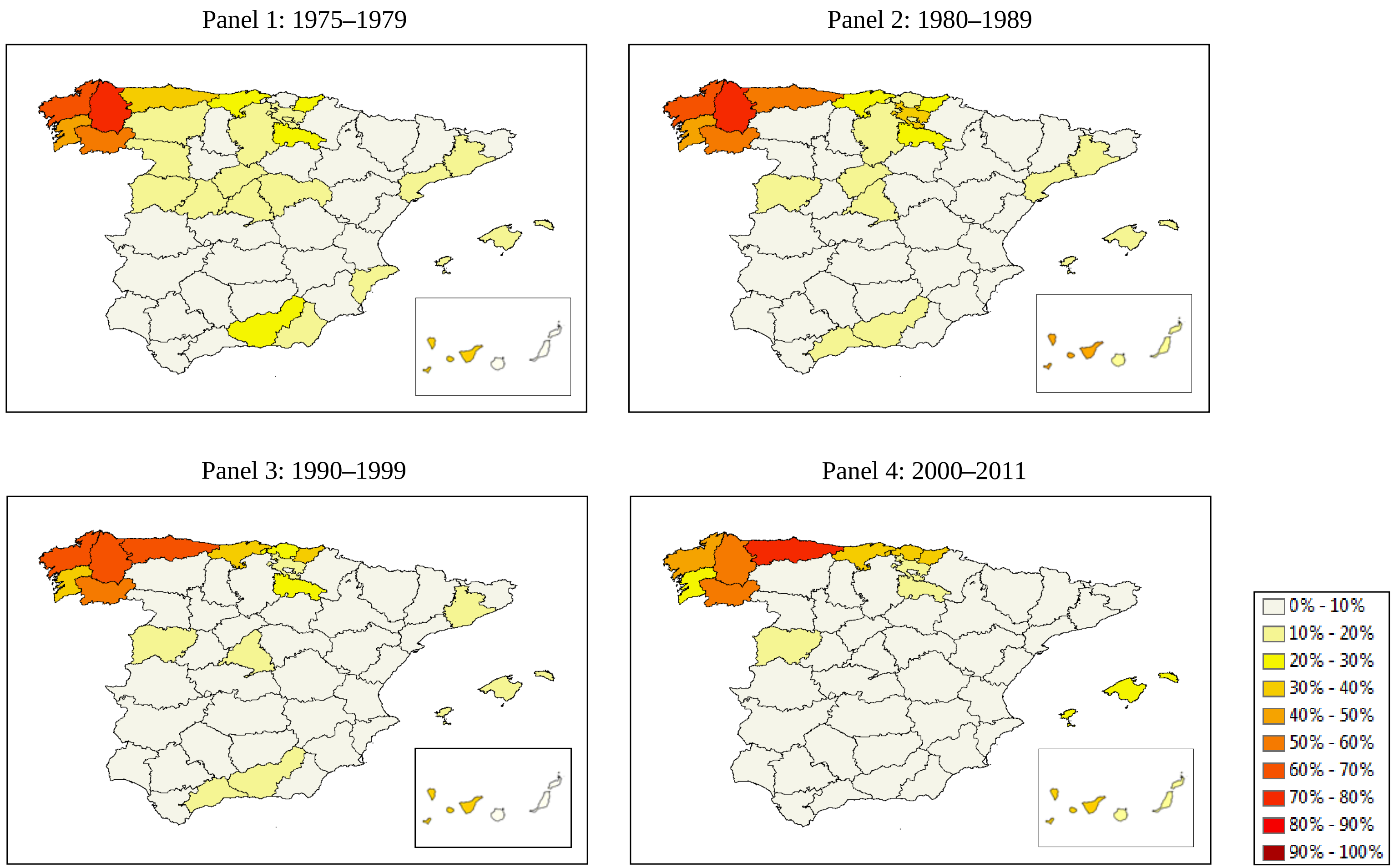
Source: AEA (MAGRAMA). Calculus: Author.

Figure 4.13: Specialization pattern by province: legumes



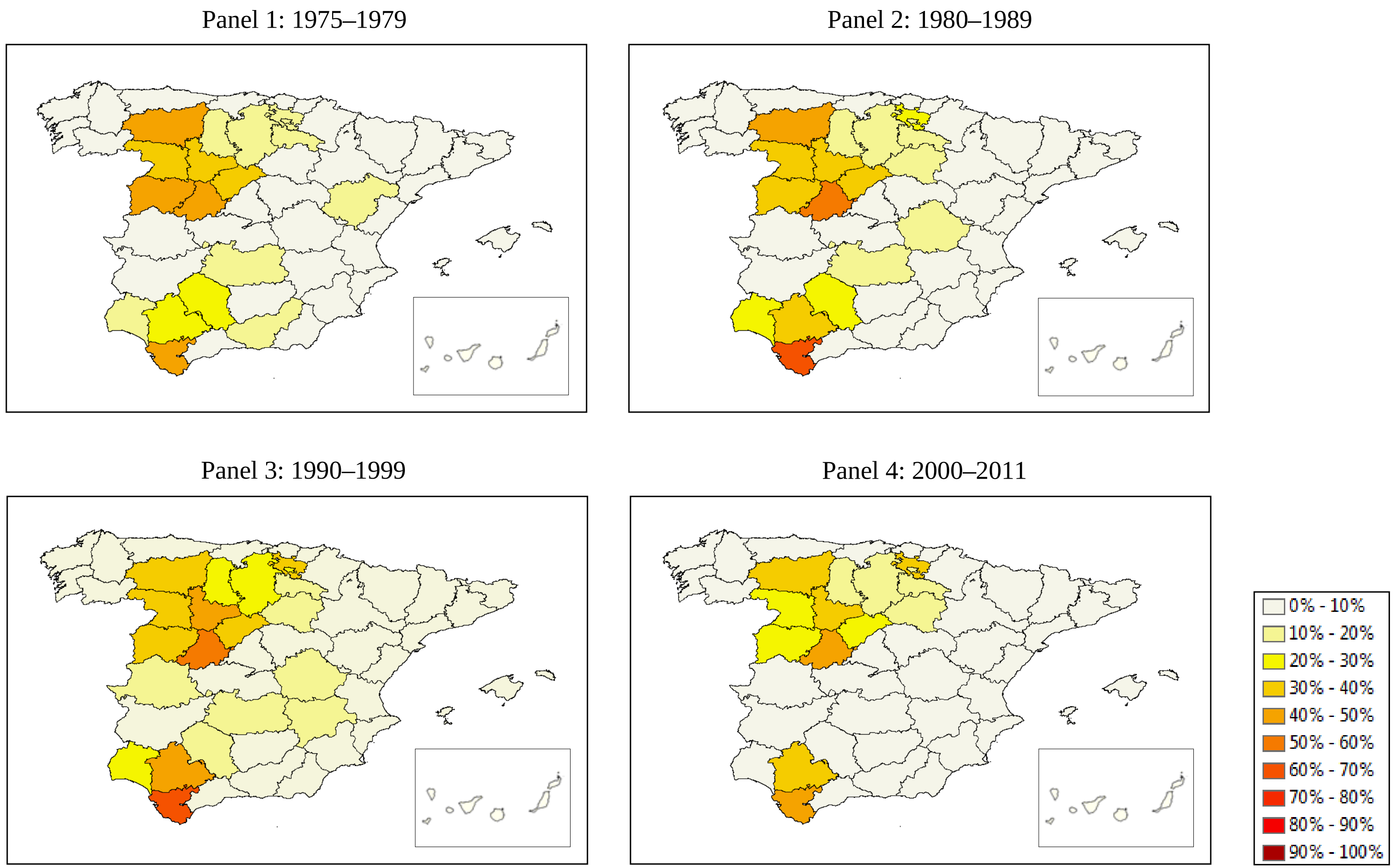
Source: AEA (MAGRAMA). Calculus: Author.

Figure 4.14: Specialization pattern by province: tubers



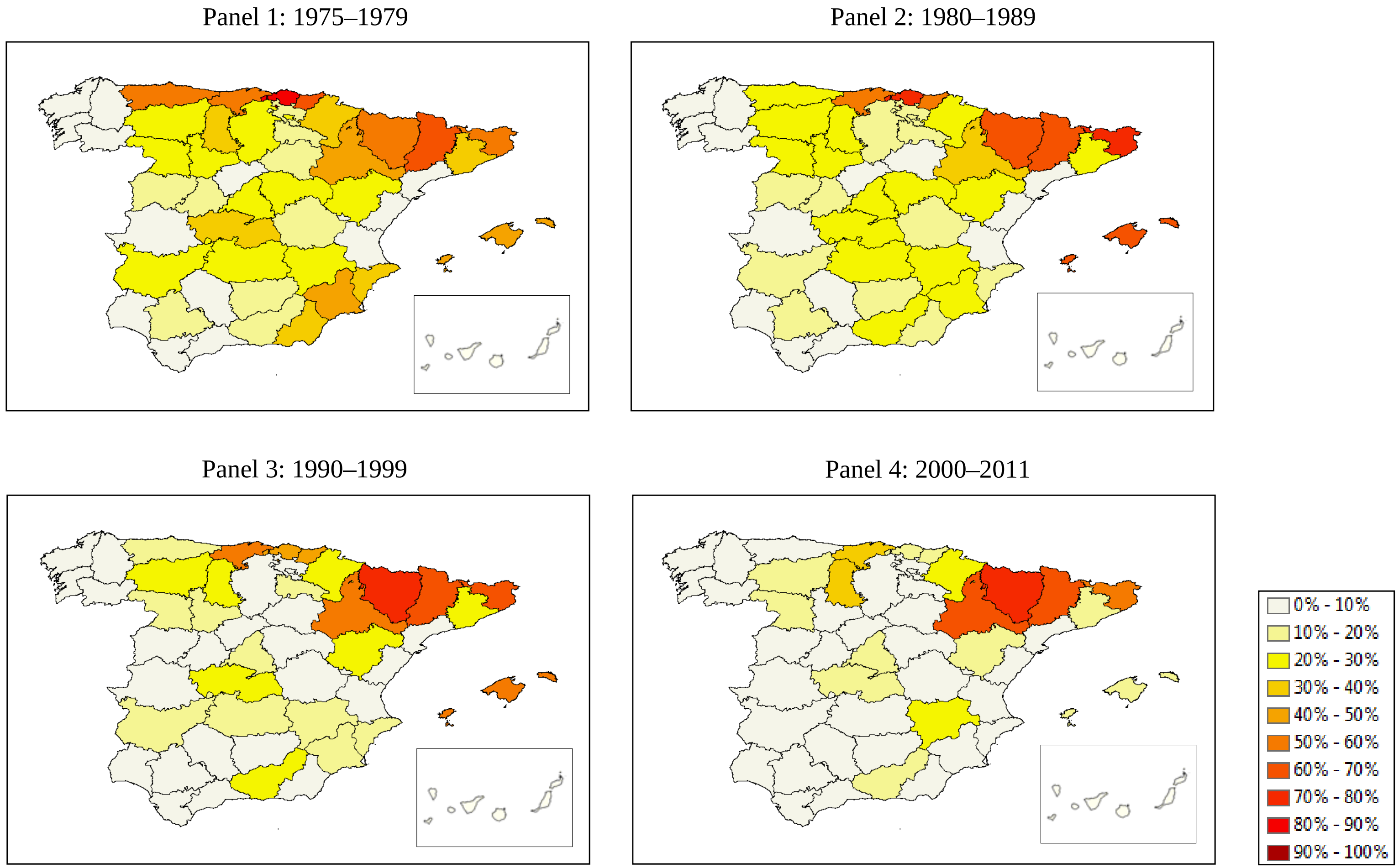
Source: AEA (MAGRAMA). Calculus: Author.

Figure 4.15: Specialization pattern by province: industrial crops



Source: AEA (MAGRAMA). Calculus: Author.

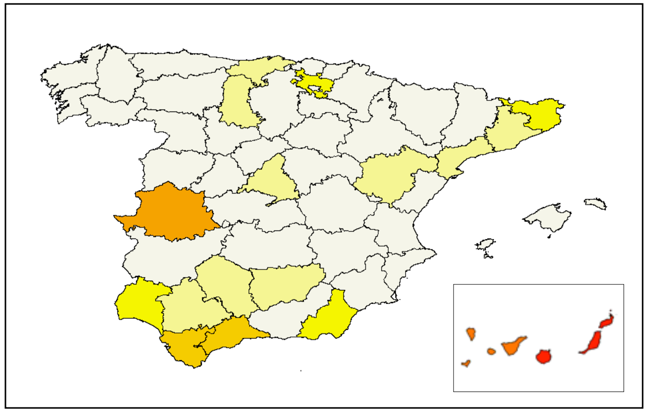
Figure 4.16: Specialization pattern by province: fodders



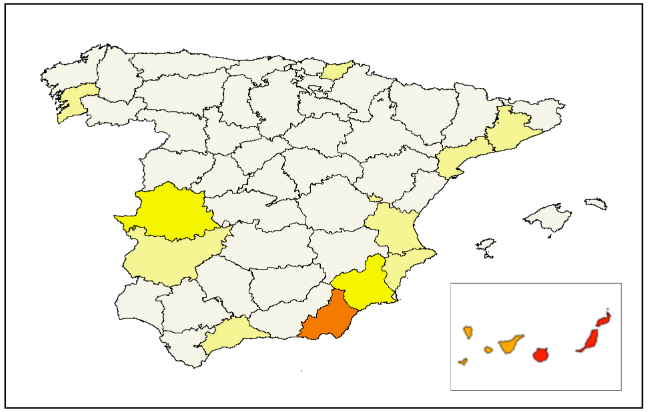
Source: AEA (MAGRAMA). Calculus: Author.

Figure 4.17: Specialization pattern by province: vegetables

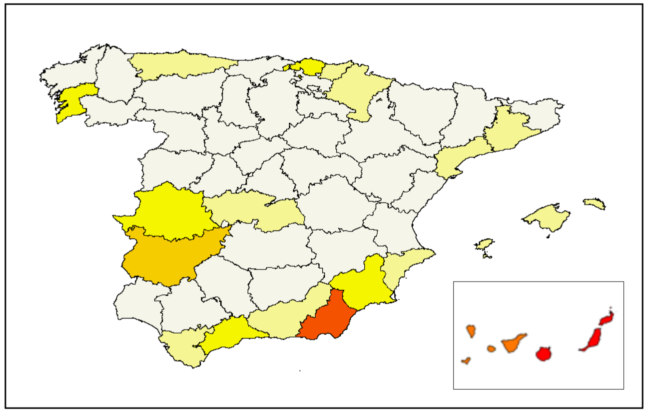
Panel 1: 1975–1979



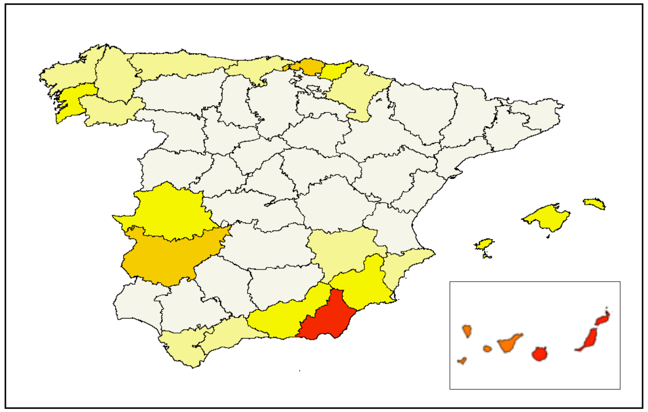
Panel 2: 1980–1989

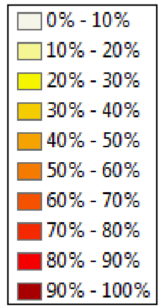


Panel 3: 1990–1999



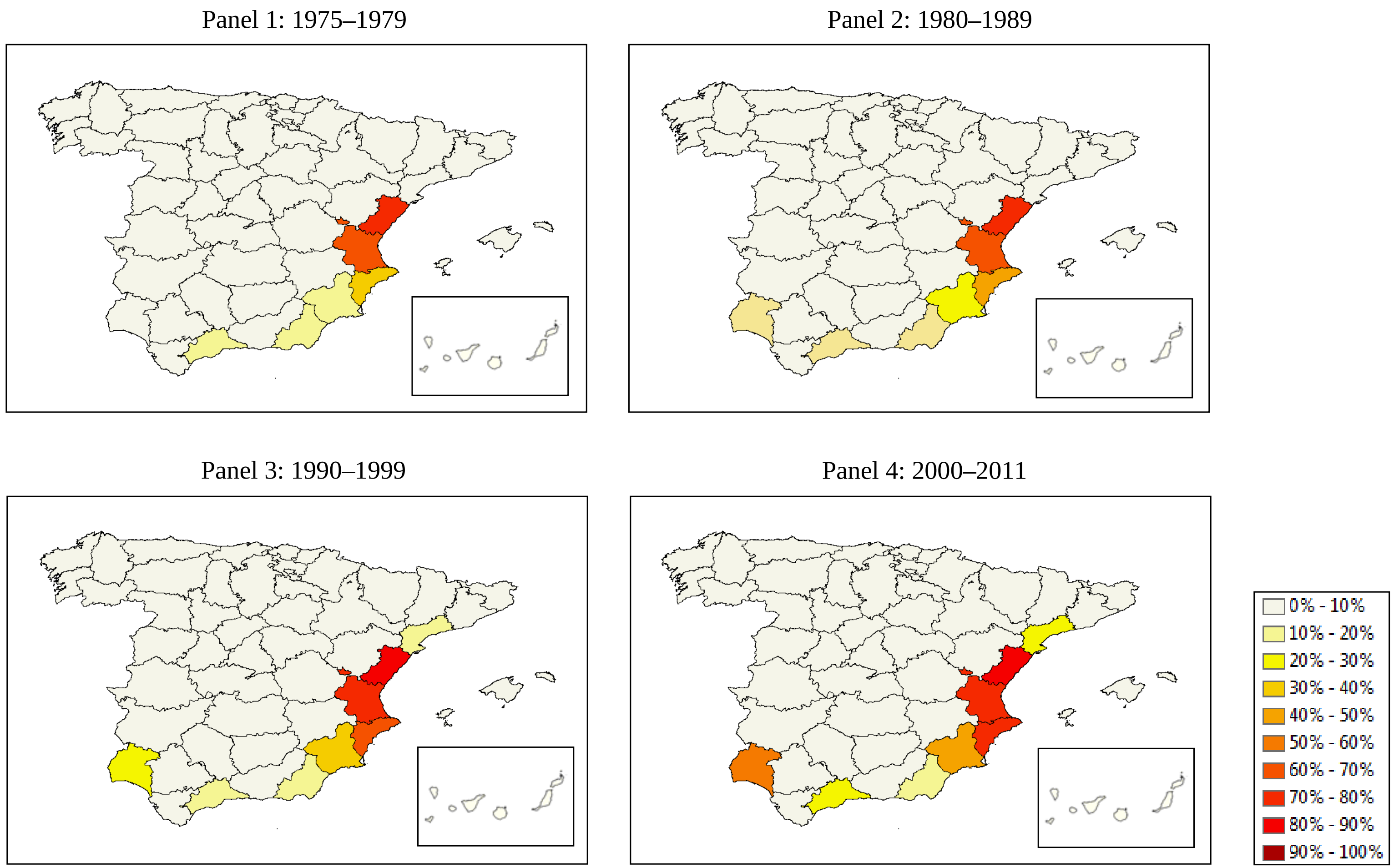
Panel 4: 2000–2011





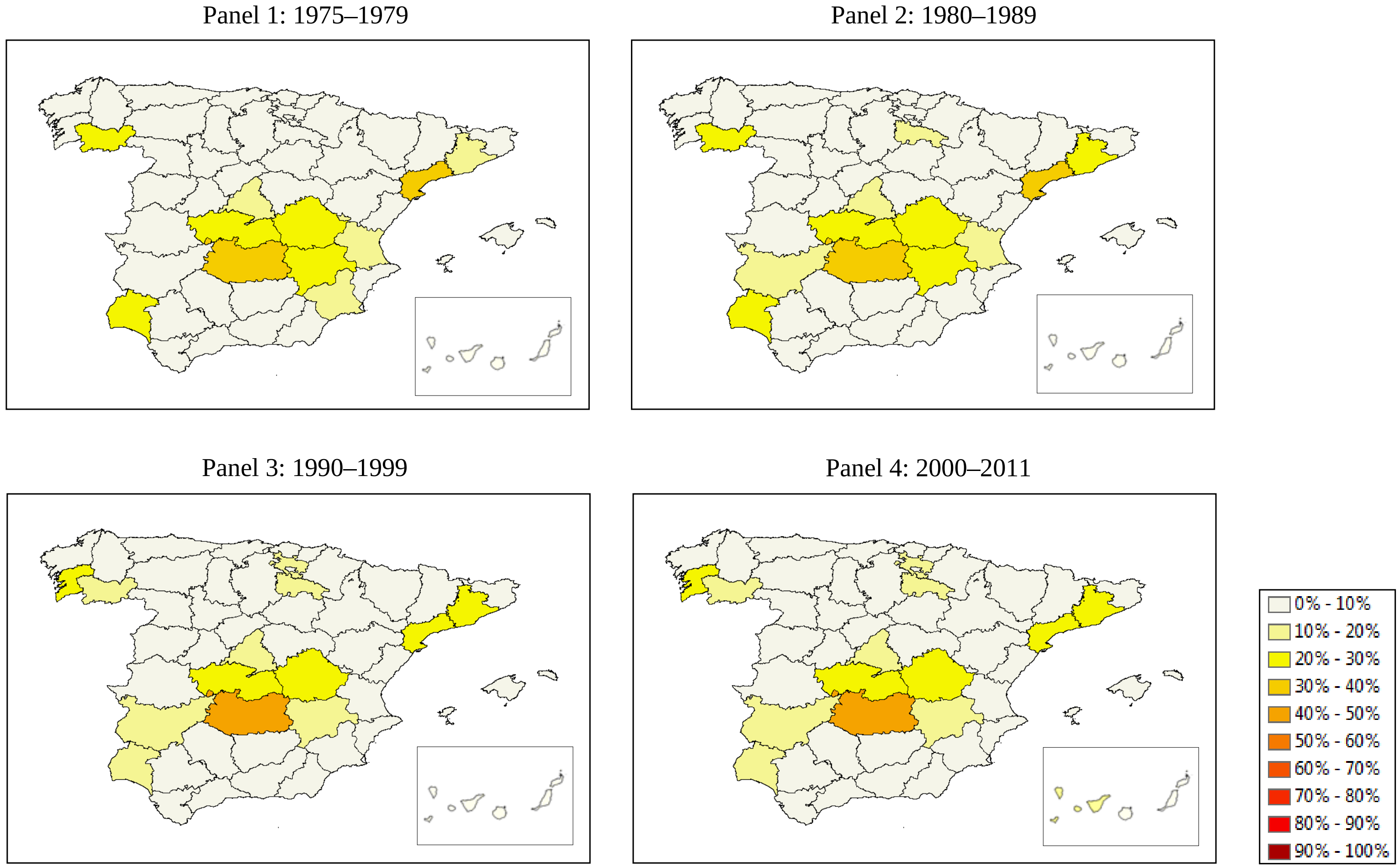
Source: AEA (MAGRAMA). Calculus: Author.

Figure 4.18: Specialization pattern by province: citrus



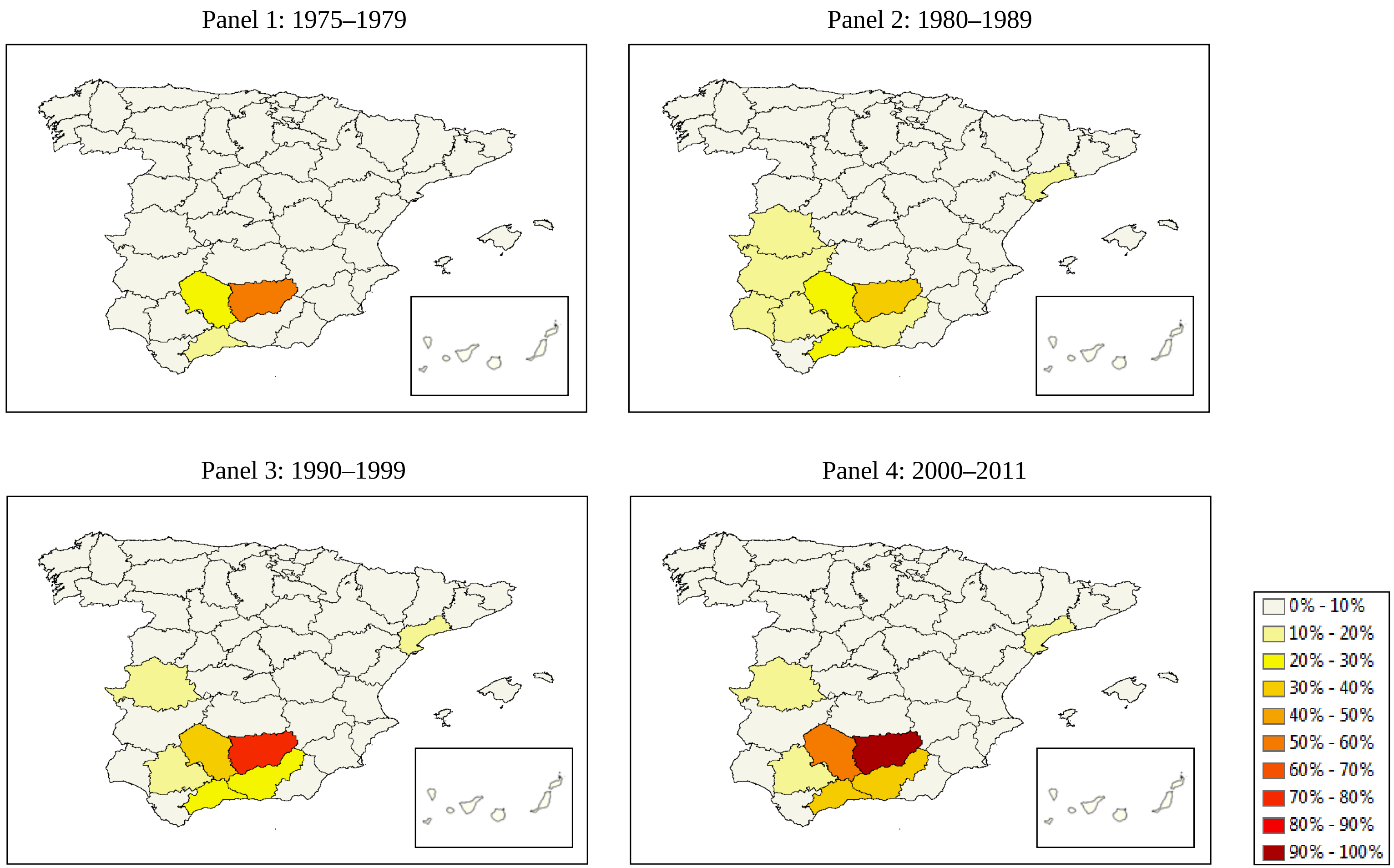
Source: AEA (MAGRAMA). Calculus: Author.

Figure 4.19: Specialization pattern by province: grapes



Source: AEA (MAGRAMA). Calculus: Author.

Figure 4.20: Specialization pattern by province: olives



Source: AEA (MAGRAMA). Calculus: Author.

Table 4.8: Herfindahl-Hirschman index, 1975 and 2011

Code	Province	HH 1975	HH 2011
1	Alava	0.1192	0.2132
2	Albacete	0.2280	0.1502
3	Alicante	0.2514	0.5122
4	Almería	0.2033	0.5571
5	Avila	0.2292	0.1829
6	Badajoz	0.1400	0.1927
7	Baleares	0.3630	0.1586
8	Barcelona	0.1664	0.2378
9	Burgos	0.1918	0.2573
10	Cáceres	0.1221	0.1995
11	Cádiz	0.2035	0.0952
12	Castellón	0.5006	0.8153
13	Ciudad Real	0.2617	0.3071
14	Córdoba	0.1598	0.4160
15	A Coruña	0.4230	0.2659
16	Cuenca	0.1895	0.3266
17	Girona	0.3327	0.3144
18	Granada	0.1334	0.2376
19	Guadalajara	0.2162	0.3198
20	Guipuzcoa	0.5022	0.1943
21	Huelva	0.1687	0.5381
22	Huesca	0.3674	0.4842
23	Jaen	0.3384	0.8956
24	León	0.2316	0.2175
25	Lleida	0.4178	0.3979
26	La Rioja	0.1537	0.2039
27	Lugo	0.4844	0.3352
28	Madrid	0.1394	0.1945
29	Málaga	0.1307	0.2760
30	Murcia	0.2803	0.2914
31	Navarra	0.2050	0.1624
32	Ourense	0.4033	0.3555
33	Asturias	0.4200	0.6239
34	Palencia	0.1885	0.2088
35	Las Palmas de Gran Canaria	0.7207	0.4094
36	Pontevedra	0.3404	0.2862
37	Salamanca	0.1956	0.1631
38	Santa Cruz de Tenerife	0.2864	0.3008
39	Cantabria	0.3440	0.1698
40	Segovia	0.2472	0.1968
41	Sevilla	0.1332	0.0868
42	Soria	0.2648	0.3449
43	Tarragona	0.1720	0.2049
44	Teruel	0.1716	0.2571
45	Toledo	0.1806	0.1825
46	Valencia	0.4165	0.6293
47	Valladolid	0.2815	0.1913
48	Vizcaya	0.7946	0.2041
49	Zamora	0.2071	0.1635
50	Zaragoza	0.2690	0.4834

Note: $0 \leq HH \leq 1$.

Chapter 5

Conclusions

Today, the meaning of the European Union stems from the commitment of all member states to unite their efforts toward peace and prosperity. Although the origin of the European integration project arose from a desire to avoid another international military conflict, it has evolved to meet new necessities and face new challenges.

The failure of the European Constitution in 2005, troubles during the ratification of the Lisbon Treaty, and the absence of effective instruments to alleviate the Eurozone crisis have together prompted discussion about the future of Europe's ambitious integration project. Nevertheless, sufficient evidence suggests that the project has clearly benefited the welfare of citizens of all member states, despite equally clear evidence that there is room to improve the effectiveness of Europe's communitarian initiatives.

This dissertation has addressed certain relevant issues related to the first steps of the European integration project. In Chapter 2, we have evaluated the extent to which the fifth EU enlargement yielded effective trade integration among EU member states, and in Chapter 3, we have emphasized determinants of the mobility of capital

As defined by Jacques Delors in July 1987, among the most valuable assets of being an EU member is the possibility of sharing a unique European market: "By its size—the biggest in the world—the single market without frontiers is an invaluable asset to revitalize our business and make them more competitive. It is one of the main engines of the European Union." (Source: Baldwin and Wyplosz, 2012, p. 165.) Our results, however, tend to underscore that not all EU members have fully benefited from the advantages of the European Union's unique market. Specifically, we have identified in Chapter 2 that the trade integration of EU-10 and EU-15 countries since 2004 has not been as thorough as expected and, in Chapter 3, that the mobility of capital and labor across member states has not been as fluid as hoped.

In Chapter 4, we have examined the European agricultural sector in terms of a controversial communitarian policy—namely, the Common Agricultural Policy (CAP)—considered to comprise a set of mixed, complex, and expensive policies with poorly defined objectives that makes its evaluation exceptionally difficult. In response, we have proposed an original empirical strategy to objectively quantify the effects of the policy in Spain. In doing so, we have identified improvement in production efficiency subsequent to the Fischler reform of 2003, which effectively broke the link between subsidies and level of output.

We have furthermore detected specific circumstances in which the effects of European integration have been different from the expected ones. In sum, if the European integration project seeks to progress, these weaknesses—most of them associated with the lack of strong economic ties among all member states—need to be overcome.

Methodological matters

This dissertation has scrutinized specific topics related to the European integration project by using original databases and updated empirical tools. Altogether, investigating why EU member states are less integrated than expected is clearly essential to rectify current weaknesses and advance the situation toward greater achievement. Of course, many other fascinating issues related to the European integration process have not been addressed here, including regional disparities and, by extension, the impacts of regional policy.

Empirical research is always conditioned by the availability and quality of data. The database built to conduct our analysis in Chapter 3 is based upon two micro-data sources: one for firms and the other for workers. Without a common identifier to match both sources of data, we had to aggregate information at the province–sector level, which entailed the loss of individual characteristics at the firm and worker levels. Accordingly, a valuable extension would involve exploiting an employer–employee database to allow us a complete micro-level analysis that considers matching a job opening with the skills of the employee filling the position. In the same spirit—and as discussed in the corresponding chapters—a potential extension of our research questions is thus to enlarge the sample of analysis by including other countries, which would allow us to make valuable international comparisons mostly overlooked in parts of this dissertation.

Another frequent problem in our empirical research relates to econometric issues and managing missing values and zeros. In Chapter 3, we sought to overcome the abundance of zeros for the dependent variable by using a count data model; however, our strategy did not perform well statistically when we accounted for fixed effects. To reduce the number of zeros, we thus excluded province–sector combinations in which no new workplaces were established during our period of study. Nevertheless, further developments need to be pursued in the line of these count models.

Lastly, the numerical exercise presented in Chapter 4 also presents a limitation. The maximization problem used to allocate crops across parcels of land does not consider any technical restriction—for instance, crop rotation and fallow periods. It would therefore be interesting to employ another algorithm that would allow us to control for these restrictions and, in turn, approximate the efficient scenario as much as possible to the real features of the agricultural sector.

Policy implications

This dissertation has highlighted a latent problem of competitiveness in some EU economies. We have identified that EU-10 countries have suffered from this problem, especially in high-tech industries, for evidence shows that after 2004, exports in some high-tech sectors increased more toward non-EU countries than toward EU-15 ones. If the competitiveness gap between EU-10 does not shrink, it would generate strong, irreversible structural differences, which would also affect the well-being of EU-10 citizens. Along those lines, European national governments should implement specific measures to manage competitiveness in their economies. For instance, they could prompt transitions toward more technology-oriented production by fueling the creation of ad-hoc structures, in which knowledge could be shared among companies, or by supporting high-tech programs to update the production process of their industrial sectors.

In Europe's transition to international integration, concerns with competitiveness take different forms. Beyond the lack of certain features such as technology, the dearth of competitiveness may also refer to a poor economic and business environment. Indeed, this dilemma emerged as a result in our study of FDI inflows in Spain. Our findings stress that Spain seems to have been unable to foster a long-term environment for attracting foreign investors, most of whom are enticed by low-cost options. As a consequence, current features of productivity in Spain are attractive to low-skilled immigrants, but not high-skilled ones who could reinforce the highly skilled workforce usually recognized to rank among the building blocks of any healthy business environment. If Spanish authorities sought to appeal to the long-term interests of foreign investors in their domestic economy, then it would be advisable to adopt the incentive scheme already consolidated in other countries that relies on positioning high-skilled local workers as vital to attracting FDI, as well as other intangible assets such as technology and an adequate business climate. Indirectly, this strategy also entails reinforcing how the educational content of training programs matches employer needs.

As a final comment regarding the CAP, despite criticism of the policy, an active and effective agricultural sector in Europe is necessary to guarantee high-quality products and preserve the environment and natural resources. However, the CAP should place fewer demands upon resources and better define its objectives. We have found that when direct payments are independent of the level of output, the market distortions produced by subsidies decrease. In the wake of this effect, it is advisable that any measure implying subsidies to farmers targets lump-sum money transfers independent of output levels and all other pecuniary conditions.

Bibliography

"Q&A: Reform and EU Farm Policy". *BBC News*. (July 1, 2013) Available online at: http://www.bbc.com/news/world-europe-11216061.

Abraham, F. and J. Konings (1999). "Does the Opening of Central and Eastern Europe Threaten Employment in the West?". *The World Economy Vol. 22*(4), pp. 585–601.

Alegría, R. (2006). *Countries, Regions and Multinational Firms: Location Determinants in the European Union.* Working Paper. Available online at: https://ideas.repec.org/p/wiw/wiwrsa/ersa06p143.html.

Amiti, M. and B. S. Javorcik (2008). "Trade Costs and Location of Foreign Firms in China". *Journal of Development Economics Vol. 85*(1–2), pp. 129–149.

Anderson, J. E. (1979). "A Theoretical Foundation for the Gravity Equation". *American Economic Review Vol. 69*(1), pp. 106–116.

Anderson, J. E. and E. van Wincoop (2003). "Gravity with Gravitas: A Solution to the Border Puzzle". *American Economic Review Vol. 93*(1), pp. 170–192.

Antimiani, A. and V. Costantini (2013). "Trade Performances and Technology in the Enlarged European Union". *Journal of Economic Studies Vol. 40*(3), pp. 355–389.

Armington, P. S. (1969). "A Theory of Demand for Products Distinguished by Place of Production". *Staff Paper IMF Vol. 16*(1), pp. 159–178.

ATA (2014). *Evolución Autónomos Extranjeros 2013.* (January 27, 2014) Federación Nacional de Asociaciones de Trabajadores Autónomos, Madrid. Available online at http://www.ata.es/publicaciones/informes.

Baier, S. L. and J. H. Bergstrand (2007). "Do Free Trade Agreements Actually Increase Members' International Trade?". *Journal of International Economics Vol. 71*(1), pp. 72–95.

Baldwin, R. and D. Taglioni (2006). *Gravity for Dummies and Dummies for Gravity Equations.* NBER Working Paper No. 12516. The National Bureau of Economic Research, Cambridge, MA. Available online at: http://www.nber.org/papers/w12516.

Baldwin, R. and C. Wyplosz (2012). *The Economics of European Integration* (4th ed.). London: McGraw-Hill Higher Education.

Baldwin, R. E., R. Forslid, and J. I. Haaland (1996). "Investment Creation and Diversion in Europe". *The World Economy Vol. 19*(6), pp. 635–659.

Baltagi, B. H. (2008). *Econometric Analysis of Panel Data* (4th ed.). Chichester: John Wiley & Sons, Ltd.

Baltagi, B. H., P. H. Egger, and M. Pfaffermayr (2003). "A Generalized Design for Bilateral Trade Flow Models". *Economics Letters Vol. 80*(3), pp. 391–397.

Barba Navaretti, G. (2004). "Host Country Effects: Conceptual Framework and the Evidence". In G. Barba Navaretti and A. J. Venables (Eds.), *Multinational Firms in the World Economy*, Chapter 7, pp. 151–185. Princenton, NJ: Princenton University Press.

Barba Navaretti, G. and A. J. Venables (2004a). "Facts and Issues". In G. Barba Navaretti and A. J. Venables (Eds.), *Multinational Firms in the World Economy*, Chapter 1, pp. 1–22. Princenton, NJ: Princenton University Press.

Barba Navaretti, G. and A. J. Venables (2004b). "The Multinational Enterprise: An Overview of Theory and Empirical Findings". In G. Barba Navaretti and A. J. Venables (Eds.), *Multinational Firms in the World Economy*, Chapter 2, pp. 23–48. Princenton, NJ: Princenton University Press.

Barciela, C., J. Giráldez, and I. López (2005). "Sector Agrario y Pesca". In A. Carreras and X. Tafunell (Eds.), *Estadística Históricas de España. Siglos XIX-XX*, Volume I, Chapter 4, pp. 245–356. Bilbao: Fundación BBVA.

Barrell, R. and N. Pain (1999). "Domestic Institutions, Agglomerations and Foreign Direct Investment in Europe". *European Economic Review Vol. 43*(4–6), pp. 925–934.

Barrios, S. and E. Strobl (2002). "Foreign Direct Investment and Productivity Spillovers: Evidence from the Spanish Experience". *Weltwirtschaftliches Archiv Vol. 138*(3), pp. 459–481.

Barry, F. G. (2004). “FDI and the Host Economy: A Case Study of Ireland”. In G. Barba Navaretti and A. J. Venables (Eds.), *Multinational Firms in the World Economy*, Chapter 8, pp. 187–216. Princenton, NJ: Princenton University Press.

Basile, R. (2004). “Acquisition versus Greenfield Investment: The Location of Foreign Manufacturers in Italy”. *Regional Science and Urban Economics Vol. 34*(1), pp. 3–25.

Bergstrand, J. H. (1985). The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence. *The Review of Economics and Statistics Vol. 67*(3), pp. 474–481.

Bhattacharya, U. and P. Groznik (2008). “Melting Pot or Salad Bowl: Some Evidence from U.S. Investments Abroad”. *Journal of Financial Markets Vol. 11*(3), pp 228–258.

Blalock, G. and P. J. Gertler (2008). “Welfare Gains from Foreign Direct Investment through Technology Transfer to Local Suppliers”. *Journal of International Economics Vol. 74*(2), pp. 402–421.

Blonigen, B. A. (2005). “A Review of the Empirical Literature on FDI Determinants”. *Atlantic Economic Journal Vol. 33*(4), pp. 383–403.

Bloom, N. and J. Van Reenen (2007). “Measuring and Explaining Management Practices across Firms and Countries”. *The Quarterly Journal of Economics Vol. 122*(4), pp. 1351–1408.

Brakman, S., H. Garretsen, and C. van Marrewijk (2011). *The New Introduction to Geographical Economics* (Reprinted with corrections, 2nd ed.). Cambridge: Cambridge University Press.

Brenton, P. A. and D. Gros (1997). “Trade Reorientation and Recovery in Transition Economies”. *Oxford Review of Economic Policy Vol. 13*(2), pp. 65–76.

Brun, J.-F., C. Carrère, P. Guillaumont, and J. de Melo (2005). “Has Distance Died? Evidence from a Panel Gravity Model”. *The World Bank Economic Review Vol. 19*(1), pp. 99–120.

Buch, C. M., J. Kleinert, and F. Toubal (2006). “Where Enterprises Lead, People Follow? Links between Migration and FDI in Germany”. *European Economic Review Vol. 50*(8), pp. 2017–2036.

Bussière, M., J. Fidrmuc, and B. Schnatz (2008). “EU Enlargement and Trade Integration: Lessons from a Gravity Model”. *Review of Development Economics Vol. 12*(3), pp. 562–576.

Bustos, P., B. Caprettini, and J. Ponticelli (2015). *Agricultural Productivity and Structural Transformation: Evidence from Brazil.* SSRN Working Paper ID 2309594. Social Science Research Network, Rochester, NY. Available online at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2309594.

BvD. *Sistema de Análisis de Balances Ibéricos (SABI).* (Accessed on April 2014) AMADEUS Bureau van Dijk Electronic Publishing (BvDEP).

Carr, D. L., J. R. Markusen, and K. E. Maskus (2001). "Estimating the Knowledge-Capital Model of the Multinational Enterprise". *American Economic Review Vol. 91*(3), pp. 693–708.

Carrère, C. (2006). "Revisiting the Effects of Regional Trade Agreements on Trade Flows with Proper Specification of the Gravity Model". *European Economic Review Vol. 50*(2), pp. 223–247.

Cebrián de Miguel, J. A., M. I. Bodega, and M. A. Martín-Lou (2007). "Inmigración y Creación de Empresas". In R. Domínguez-Martí (Ed.), *Inmigración: Crecimiento Económico e Integración Social*, pp. 147–166. Santander: Universidad de Cantabria y Parlamento de Cantabria.

Chakravarty, S. R. (2005). *Microeconomics.* New Delhi: Allied Published Pvt. Limited.

Chaney, T. (2008). "Distorted Gravity: The Intensive and Extensive Margins of International Trade". *American Economic Review Vol. 98*(4), pp. 1707–1721.

Chaney, T. (2013). *The Gravity Equation in International Trade: An Explanation.* NBER Working Paper No. 19285. The National Bureau of Economic Research, Cambridge, MA. Available online at: http://www.nber.org/papers/w19285.

Chen, N. (2004). "Intra-National versus International Trade in the European Union: Why do National Borders Matter?". *Journal of International Economics Vol. 63*(1), pp. 93–118.

Coe, D. T., A. Subramanian, and N. T. Tamirisa (2007). *The Missing Globalization Puzzle: Evidence of the Declining Importance of Distance.* IMF Working Paper No. 07/126. International Monetary Fund, Washington, D.C. Available online at: http://www.imf.org/external/pubs/ft/staffp/2007/01/coe.htm.

Combes, P.-Ph., M. Lafourcade, and T. Mayer (2005). "The Trade-Creating Effects of Business and Social Networks: Evidence from France". *Journal of International Economics Vol. 66*(1), pp. 1–29.

Combes, P.-Ph., T. Mayer, and J.-F. Thisse (2008). *Economic Geography: The Integration of Regions and Nations*. Princeton, NJ: Princenton University Press.

Costinot, A. (2009). "An Elementary Theory of Comparative Advantage". *Econometrica Vol. 77*(4), pp. 1165–1192.

Costinot, A. and D. Donaldson (2012). "Ricardo's Theory of Comparative Advantage: Old Idea, New Evidence". *American Economic Review: Papers and Proceedings Vol. 102*(3), pp. 453–458.

Crozet, M., T. Mayer, and J.-L. Mucchielli (2004). "How Do Firms Agglomerate? A Study of FDI in France". *Regional Science and Urban Economics Vol. 34*(1), pp. 27–54.

de la Mata, T. and C. Llano (2013). "Social Networks and Trade of Services: Modelling Interregional Flows with Spatial and Network Autocorrelation Effects". *Journal of Geographical Systems Vol. 15*(3), pp. 319–367.

de la Rica, S., A. Glitz, and F. Ortega (2014). "Immigration in Europe: Trends, Policies, and Empirical Evidence". In B. R. Chiswick and P. W. Miller (Eds.), *Handbook of the Economics of International Migration*, Volume 1, Chapter 24, pp. 1303–1362. North-Holland: Elsevier.

Disdier, A.-C. and T. Mayer (2004). "How Different Is Eastern Europe? Structure and Determinants of Location Choices by French Firms in Eastern and Western Europe". *Journal of Comparative Economics Vol. 32*(2), pp. 280–296.

Docquier, F., Ç. Özden, and G. Peri (2014). "The Labour Market Effects of Immigration and Emigration in OECD Countries". *The Economic Journal Vol. 124*(579), pp. 1106–1145.

Donald, P. F., G. Pisano, M. D. Rayment, and D. J. Pain (2002). "The Common Agricultural Policy, EU Enlargement and the Conservation of Europe's Farmland Birds". *Agriculture, Ecosystems & Environment Vol. 89*(3), pp. 167–182.

Duch, N., D. Montolio, and M. Mediavilla (2009). "Evaluating the Impact of Public Subsidies on a Firm's Performance: A Two-Stage Quasi-Experimental Approach". *Investigaciones Regionales Vol. 16*, pp. 143–165.

Dunning, J. H. (1977). "Trade, Location of Economic Activity and the MNE: A Search for an Eclectric Approach". In B. Ohlin, P. O. Hesselborn, and P. M. Wijkman (Eds.), *The International Allocation of Economic Activity*, pp. 395–418. London: Macmillan.

Dunning, J. H. (1981). *International Production and the Multinational Enterprise*. London: George Allen & Unwin.

EC (1995). *CORINE Land Cover*. Model Documentation. European Commission, Brussels. Available online at: http://www.eea.europa.eu/publications/COR0-landcover.

EC (2001a). Enlargement of the European Union: An Historic Opportunity. Enlargement Directorate General, European Commission, Brussels. Available online at: http://ec.europa.eu/enlargement/archives/pdf/press_corner/publications/corpus_en.pdf.

EC (2001b, 2012a). *Agriculture in the European Union. Statistical and Economic Information*. Directorate-General for Agriculture and Rural Development, European Commission, Brussels. Available online at: http://ec.europa.eu/agriculture/statistics/agricultural/index_en.htm.

EC (2012b). *The Common Agricultural Policy. A Story to Be Continued*. Directorate-General for Agriculture and Rural Development, European Commission, Brussels. Available online at: http://ec.europa.eu/agriculture/cap-history/index_en.htm.

EEA. *Data and maps*. (Accessed on July 2014) European Environment Agency, Copenhagen. Available online at: http://www.eea.europa.eu/data-and-maps.

Egger, P. H. and M. Larch (2011). "An Assessment of the Europe Agreements' Effects on Bilateral Trade, GDP, and Welfare". *European Economic Review Vol. 55*(2), pp. 263–279.

Egger, P. H. and M. Pfaffermayr (2003). "The Proper Panel Econometric Specification of the Gravity Equation: A Three-Way Model with Bilateral Interaction Effects". *Empirical Economics Vol. 28*(3), pp. 571–580.

Ekholm, K. and K. H. Midelfart-Knarvic (2004). "Determinants of FDI: The Evidence". In G. Barba Navaretti and A. J. Venables (Eds.), *Multinational Firms in the World Economy*, Chapter 6, pp. 127–150. Princenton, NJ: Princenton University Press.

Espasa (1998). "Olivo y Vid, Plantas Mediterráneas". In *Temática Multimedia QUORUM*, Volume 2, p.403. Espasa-Calpe, Barcelona: Grupo Planeta.

ESRI (2011). *ArcGIS Desktop: Release 10*. Rendlands, CA: Environmental Systems Research Institute.

Eurostat. *EU Direct Investments - Main Indicators (bop_fdi_main)*. (Accessed on June 2014) Eurostat, European Commission, Luxembourg. Available online at:

http://ec.europa.eu/eurostat/web/main [Data / Database / Database by theme / Economy and finance / Balance of payments - International transactions / European Union direct investments].

Eurostat. *EU Trade since 1988 by SITC (DS-018995).* (Accessed on March 2013) Eurostat, European Commission, Luxembourg. Available online at: http://ec.europa.eu/eurostat/web/main [Data / Database / Database by theme / International trade / International trade detailed data (detail)].

Eurostat. *Euro/ECU Exchange Rates - Annual Data (ert_bil_eur_a).* (Accessed on January 2013) Eurostat, European Commission, Luxembourg. Available online at: http://ec.europa.eu/eurostat/web/main [Data / Database / Database by theme / Economy and finance / Exchanges rates / Bilateral exchanges rates / Euro/ECU exchange rates].

Eurostat. *Former Euro Area National Currrencies vs. euro/ECU - Annual Data (ert_h_eur_a).* (Accessed on August 2014) Eurostat, European Commission, Luxembourg. Available online at: http://ec.europa.eu/eurostat/web/main [Data / Database / Database by theme / Economy and finance / Exchanges rates / Exchanges rates: historical data / Former euro area national currencies vs. euro/ECU].

Eurostat (2012). *Compact guides: European Neighbourhood Policy - South Countries. International Trade in Goods.* (December 21, 2012) Collection: Compact Guides and Catalogues. Theme: International Trade. Product code: KS-32-12-269. Eurostat, European Commission, Luxembourg. Available online at: http://ec.europa.eu/eurostat/web/main [Publications / Collections].

Eurostat (2015a). *Unemployment Statistics.* Statistics Explained. Eurostat, European Commission, Luxembourg. Available online at: http://ec.europa.eu/eurostat/statistics-explained/index.php/Main_Page.

Eurostat (2015b). *Agricultural Census in Spain.* Statistics Explained. Eurostat, European Commission, Luxembourg. Eurostat, European Commission, Luxembourg. Available online at: http://ec.europa.eu/eurostat/statistics-explained/index.php/Main_Page.

FAOSTAT. *Producer Prices.* (Accessed on September 2014) Statistics Division, Food and Agriculture Organization of the United Nations, Rome. Available online at: http://faostat3.fao.org/home/E [Download data / Prices].

Faucompret, E. Konings, J. and H. Vandenbussche (1999). "The Integration of Central and Eastern Europe in the European Union: Trade and Labour Market Adjustment". *Journal of World Trade Vol. 33*(6), pp. 121–145.

FBBVA and Ivie (2014). *El Stock y los Servicios del Capital en España y su Distribución Territorial y Sectorial (1964-2012).* (Accessed on May 2015) Fundación BBVA and Instituto Valenciano de Investigaciones Económicas. Available online at: http://www.fbbva.es/TLFU/microsites/stock09/fbbva_stock08_index.html.

Fearne, A. (1997). "The History and Development of the CAP 1945-1990". In C. Ritson and D. R. Harvey (Eds.), *The Common Agricultural Policy*, Chapter 2, pp. 11–55. Wallingford: CAB INTERNATIONAL.

Feenstra, R. C. (2004). *Advanced International Trade: Theory and Evidence.* Princeton, NJ: Princeton University Press.

Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015). "The Next Generation of the Penn World Table". *American Economic Review Vol. 105*(10), pp. 3150–3182.

Felbermayr, G. J. and W. Kohler (2006). "Exploiting the Intensive and Extensive Margins of World Trade". *Review of World Economics Vol. 142*(4), pp. 642–674.

Ferreiro, J., C. Rodriguez, and F. Serrano (1997). "The Role of Foreign Direct Investment in an Old Industrial Region: The Case of the Basque Country and Japanese FDI". *European Planning Studies Vol. 5*(5), pp. 637–657.

Figueiredo, O., P. Guimarães, and D. Woodward (2002). "Home-Field Advantage: Location Decisions of Portuguese Entrepreneurs". *Journal of Urban Economics Vol. 52*(2), pp. 341–361.

Foad, H. (2012). "FDI and Immigration: A Regional Analysis". *The Annals of Regional Science Vol. 49*(1), pp. 237–259.

Fontagné, L., M. Freudenberg, and M. Pajot (1999). "Le Potentiel d'Échanges entre l'Union Européenne et les PECO. Un Réexamen". *Revue Économique Vol. 50*(6), pp. 1139–1168.

Frankel, J. A. (1997). *Regional Trading Blocs In the World Economic System.* Washington, D.C.: Peterson Institute for International Economics.

Fundación-Bancaja and Ivie (2014). *Capital Humano en España y su Distribución Provincial.* (Accessed on May 2015) Fundación Bancaja and Instituto Valenciano de Investigaciones Económicas. Available online at: http://www.ivie.es/es/banco/caphum/series.php.

Gao, T. (2003). "Ethnic Chinese Networks and International Investment: Evidence from Inward FDI in China". *Journal of Asian Economics Vol. 14*(4), pp. 611–629.

Gleditsch, K. S. (2002). "Expanded Trade and GDP Data". *Journal of Conflict Resolution Vol. 46*(5), pp. 712–724.

Glick, R. and A. K. Rose (2002). "Does a Currency Union Affect Trade? The Time-Series Evidence". *European Economic Review Vol. 46*(6), pp. 1125–1151.

González, L. and F. Ortega (2011). "How do Very Open Economies Adjust Large Immigration Flows? Recent Evidence from Spanish Regions". *Labour Economics Vol. 18*(1), pp. 57–70.

Gould, D. M. (1994). "Immigrant Links to the Home Country: Empirical Implications for U.S. Bilateral Trade Flows". *The Review of Economics and Statistics Vol. 76*(2), pp. 302–316.

Griffith, R., S. Redding, and H. Simpson (2003). *Productivity Convergence and Foreign Ownership at the Establishment Level.* CEP Discussion Paper 573. Center for Economic Performance, London. Available online at: http://cep.lse.ac.uk/_new/publications/.

Griffith, R., S. Redding, and H. Simpson (2004). *Foreign Ownership and Productivity: New Evidence from the Service Sector and the R&D Lab.* CEP Discussion Paper 649. Center for Economic Performance, London. Available online at: http://cep.lse.ac.uk/_new/publications/.

Gros, D. and A. Gonciarz (1996). "A Note of the Trade Potential of Central and Eastern Europe". *European Journal of Political Economy Vol. 12*(4), pp. 709–721.

Gros, D. and A. Steinherr (1995). *Winds of Change: Economic Transition in Central and Eastern Europe.* Harlow: Longman.

Guimarães, P., O. Figueiredo, and D. Woodward (2000). "Agglomeration and the Location of Foreign Direct Investment in Portugal". *Journal of Urban Economics Vol. 47*(1), pp. 115–135.

Guimarães, P. and P. Portugal (2010). "A Simple Feasible Procedure to Fit Models with High-Dimensional Fixed Effects". *The Stata Journal Vol. 10*(4), pp. 628–649.

Halvorsen, R. and R. Palmquist (1980). "The Interpretation of Dummy Variables in Semilogarithmic Equations". *American Economic Review Vol. 70*(3), pp. 474–475.

Harris, M. N., L. Kónya, and L. Mátyás (2012). "Some Stylized Facts about International Trade Flows". *Review of International Economics Vol. 20*(4), 781–792.

Haskel, J. E., S. C. Pereira, and M. J. Slaughter (2002). *Does Inward Foreign Direct Investment Boost the Productivity of Domestic Firms?* NBER Working Paper No. 8724. The National Bureau of Economic Research, Cambridge, MA. Avaialable online at: http://www.nber.org/papers/w8724.

Head, K. and T. Mayer (2000). "Non-Europe: The Magnitude and Causes of Market Fragmentation in the EU". *Weltwirtschaftliches Archiv Vol. 136*(2), pp. 284–314.

Head, K. and T. Mayer (2004). "The Empirics of Agglomeration and Trade". In J. V. Henderson and J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Volume 4, Chapter 59, pp. 2609–2669. Amsterdam: Elsevier.

Head, K., T. Mayer, and J. Ries (2010). "The Erosion of Colonial Trade Linkages after Independence". *Jounal of International Economics Vol. 81*(1), pp. 1–14.

Head, K. and J. Ries (1998). "Immigration and Trade Creation: Econometric Evidence from Canada". *Canadian Journal of Economics Vol. 31*(1), pp. 47–62.

Head, K., J. Ries, and D. Swenson (1995). "Agglomeration Benefits and Location Choice: Evidence from Japanese Manufacturing Investments in the United States". *Journal of International Economics Vol. 38*(3–4), pp. 223–247.

Helpman, E., M. J. Melitz, and S. R. Yeaple (2004). Exports versus FDI with Heterogeneous Firms. *American Economic Review Vol. 94*(1), pp. 300–316.

Herderschee, H. and Z. Qiao (2007). *Impact of Intra-European Trade Agreements, 1990-2005: Policy Implications for the Western Balkans and Ukraine.* IMF Working Paper No. 07/126. International Monetary Fund, Washington, D.C. Available online at: http://www.imf.org/external/pubs/cat/longres.aspx?sk=20974.0.

Hornok, C. (2010). "Trade-Enhancing EU Enlargement and the Resurgence of East-East Trade. *Focus on European Economic Integration Issue 3*, pp. 79–94.

IIASA/FAO (2012a). *Global Agro-ecological Zones (GAEZ v3.0).* (Accessed on October 2013) IIASA, Laxenburg and FAO, Rome. Available online at: http://webarchive.iiasa.ac.at/Research/LUC/GAEZv3.0/.

IIASA/FAO (2012b). *Global Agro-ecological Zones (GAEZ ver 3.0). Model Documentation.* IIASA, Laxenburg and FAO, Rome. Available online at: http://webarchive.iiasa.ac.at/Research/LUC/GAEZv3.0/.

IMF (2009). *Balace of Payments and International Investment Position Manual* (6th ed.). International Monetary Found, Washington, D.C.

IMF (2012). *World Economic Outlook Database.* (October 2012) International Monetary Found, Washington, DC. Available online at: http://www.imf.org/external/index.htm [Data / World Economic Outlook Databases].

INE. *Cifras de Población.* (Accessed on May 2015) INEbase, Instituto Nacional de Estadística, Madrid. Available online at: http://www.ine.es [INEbase / Demografía / Cifras de población y Censos demográficos].

INE. *Contabilidad Regional de España.* (Accessed on May 2015) INEbase, Instituto Naciona de Estadística, Madrid. Available online at: http://www.ine.es [INEbase / Economía / Cuentas Económicas].

Javorcik, B. S. (2004). "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages". *American Economic Review Vol. 94* (3), pp. 605–627.

Javorcik, B. S., Ç. Özden, M. Spatareanu, and C. Neagu (2011). "Migration Networks and Foreign Direct Investment". *Journal of Development Economics Vol. 94* (2), pp. 231–241.

Jeffery, S. (2003). "The EU Common Agricultural Policy". *The Guardian.* (June 26, 2013) Available online at: http://www.theguardian.com/world/2003/jun/26/eu.politics1.

Jongeneel, R. and H. Brand (2010). "Direct Income Support and Cross-Compliance". In A. Oskam, G. Meester, and H. Silvis (Eds.), *EU Policy, for Agriculture, Food and Rural Areas*, Chapter 9, pp. 191–205. Wageningen: Wageningen Academic Publishers.

Kemeny, T. (2010). "Does Foreign Direct Investment Drive Technological Upgrading?". *World Development Vol. 38* (11), pp. 1543–1554.

King, G., J. Honaker, A. Joseph, and K. Scheve (2001). "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation". *American Political Science Review Vol. 95* ((1)), pp. 49–69.

Konings, J. (2001). "The Effects of Foreign Direct Investment on Domestic Firms. Evidence from Firm-Level Panel Data in Emerging Economies". *Economics of Transition Vol. 9*(3), pp. 619–633.

Kugler, M. and H. Rapoport (2007). "International Labor and Capital Flowns: Complements or Substitutes?". *Economics Letters Vol. 94*(2), pp. 155–162.

Lafourcade, M. and E. Paluzie (2011). "European Integration, Foreign Direct Investment (FDI), and the Geography of French Trade". *Regional Studies Vol. 45*(4), pp. 419–439.

Lipsey, R. E. (2006). *Measuring the Impacts of FDI in Central and Eastern Europe.* NBER Working Paper No. 12808. The National Bureau of Economic Research, Cambridge, MA. Available online at: http://www.nber.org/papers/w12808.

Lissarrague García-Gutiérrez, J. R. and F. Martínez de Toda Fernández (2009). *Informe Vitícola. Cuestiones Referentes al Sector del Viñedo más Relevantes para la Definición de la Política de Seguros Agrarios: Situación Actual y Tendencias a Corto y Medio Plazo.* Solicitado por la Entidad Estatal de Seguros Agrarios. Available online at: http://aplicaciones.magrama.es/documentos_pwe/seminarios/vinedo_upm_ur.pdf.

MAGRAMA. *Anuario de Estadística Agraria.* (Accessed on December 2013) Ministerio de Agricultura Alimentación y Medio Ambiente, Madrid. Available online at: http://www.magrama.gob.es/es/estadistica/temas/publicaciones/anuario-de-estadistica.

MAGRAMA. *Núcleos Urbanos.* (Accessed on October 2013) Sistema Integrado de Información de Agua, Ministerio de Agricultura, Alimentación y Medio Ambiente, Madrid. Available online at: http://servicios2.marm.es/sia/visualizacion/descargas/mapas.jsp.

Markusen, J. R. (2002). *Multinational Firms and the Theory of International Trade.* Cambridge, MA: The MIT Press.

Markusen, J. R. and K. E. Maskus (2001). "Multinational Firms: Reconciling Theory and Evidence," NBER chapters. In M. Blomström and L. S. Goldberg (Eds.), *Topics in International Economics. A Festschrift in Honor of Robert E. Lipsey*, Chapter 3, pp. 71–95. Cambridge, MA: National Bureau of Economic Research, Inc.

Markusen, J. R. and K. E. Maskus (2002). "Discriminating among Alternative Theories of Multinational Enterprise". *Review of International Economics Vol. 10*(4), pp. 694–707.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic Theory.* Oxford: Oxford University Press, Inc.

Mátyás, L. (1997). "Proper Econometric Specification of the Gravity Model". *The World Economy Vol. 20*(3), pp. 363–368.

Mayer, T. and S. Zignago (2011). Notes on CEPII's Distances Measures: the GeoDist Database. CEPII Working Paper 2011-25. http://www.cepii.fr/anglaisgraph/bdd/distances.htm.

McCallum, J. (1995). "National Borders Matter: Canada-U.S. Regional Trade Patterns". *American Economic Review Vol. 85*(3), pp. 615–623.

Midelfart-Kanarvic, K. H., H. G. Overman, S. J. Redding, and A. J. Venables (2000). *The Location of European Industry.* Economic Papers No. 142. European Commission, Brussels. Available online at: http://ec.europa.eu/economy_finance/publications/index_en.htm.

MINECO. *DataComex. Estadísticas del Comercio Exterior.* (Accessed on May 2015) Ministerio de Economía y Competitividad, Madrid. Available online at: http://datacomex.comercio.es/index.htm.

Morales, E., G. Sheu, and A. Zahler (2014). *Extended Gravity.* Working paper. https://sites.google.com/site/edumoralescasado/home/research.

Myro, R. and C. M. Fernández-Otheo (2008). "Deslocalización de Empresas e Inversión Extranjera Directa en la Industria Española". *Papeles de Economía Española No. 16*, 184–202.

Nitsch, V. (2000). "National Borders and International Trade: Evidence from the European Union". *Canadian Journal of Economics Vol. 33*(4), pp. 1091–1105.

Noël, P. (2008). *Beyond Dependence: How to Deal with Russia Gas.* ECFR/09. European Council of Foreign Relations, London. Available online at: http://www.ecfr.eu/publications (Available online January 9, 2009).

Nunn, N. and N. Qian (2011). "The Potato's Contribution to Population and Urbanization: Evidence from a Historical Experiment". *The Quarterly Journal of Economics Vol. 126*(2), pp. 593–650.

Oñate, J. J., I. Atance, I. Bardají, and D. Llusia (2007). "Modelling the Effects of Alternative CAP Policies for the Spanish High-Nature Value Cereal-Steppe Farming Systems". *Agricultural Systems Vol. 94*(2), pp. 247–260.

OECD (2004). *Analysis on the 2003 CAP reform.* OECD Publishing, Paris.

OECD (2008). *OECD Benchmark Definition of Foreign Direct Investment* (4th ed.). OECD Publishing, Paris.

OECD (2012). *Agricultural Policy Monitoring and Evaluation 2012: OECD Countries.* OECD Publishing, Paris. Available online at: http://www.oecd-ilibrary.org/agriculture-and-food/agricultural-policy-monitoring-and-evaluation-2012_agr_pol-2012-en.

OECD (2015a). *Adult Education Level (Indicator).* doi:10.1787/36bce3fe-en. (Accessed on September 28, 2015) Available online at: https://data.oecd.org/.

OECD (2015b). *OECD Skills Strategy Diagnostic Report. Spain 2015.* OECD Publishing, Paris.

Paas, T. (2003). *Regional Integration and International Trade in the Context of EU Eastward Enlargement.* HWWA Discussion Paper No. 218. Hamburg Institute of International Economics, Hamburg. Available online at: https://ideas.repec.org/s/zbw/hwwadp.html.

Paniagua-Mazorra, A. (2001). "Agri-Environmental Policy in Spain. The Agenda of Socio-Political Developments at the National, Regional and Local Levels". *Journal of Rural Studies Vol. 17*, pp. 81–97.

Pantulu, J. and J. P. H. Poon (2003). "Foreign Direct Investment and International Trade: Evidence from the US and Japan". *Journal of Economic Geography Vol. 3*, pp. 241–259.

Passerini, P. (2001). *Bulk of EU Direct Investment Made between Member States in 2000.* (July 5, 2001) Collection: Statistics in Focus. Theme: Economy and Finance. Product code: KS-NJ-01-020. Eurostat, European Commission, Luxembourg. Available online at: http://ec.europa.eu/eurostat/web/main [Publications / Collections].

Rauch, J. E. (2001). "Business and Social Networks in International Trade". *Journal of Economics Literature Vol. 39*(4), pp. 1177–1203.

Rettman, A. (2007). "EU and Russia Tackle Thorny Issues at Samara Summit". *Euobserver*. Available online at: http://euobserver.com/foreign/24088.

Ritson, C. (1997). "Introduction". In A. Oskam, G. Meester, and H. Silvis (Eds.), *The Common Agricultural Policy*, Chapter 1, pp. 1–8. Wallingford: CAB INTERNATIONAL.

Rose, A. K. and E. van Wincoop (2001). "National Money as a Barrier to International Trade: The Real Case for Currency Union". *American Economic Review Vol. 91*(2), pp. 386–390.

Samuelson, P. A. (1951). "Abstract of a Theorem Concerning Substitutability in Open Leontief Models". In T. C. Koopmans (Ed.), *Activity Analysis of Production and Allocation*, pp. 142–146. New York: John Wiley & Sons.

Samuelson, P. A. (1954). "The Transfer Problem and Transport Costs, II: Analysis of Effects of Trade Impediments". *The Economic Journal Vol. 64*(254), pp. 264–289.

Santos Silva, J. M. C. and S. Tenreyro (2006). "The log of the Gravity". *The Review of Economics and Statistics Vol. 88*(4), pp. 641–658.

Santos Silva, J. M. C. and S. Tenreyro (2011). "poisson: Some Convergence Issue". *The Stata Journal Vol. 11*(2), pp. 207–212.

Sembenelli, A. and G. Siotis (2008). " Foreign Direct Investment and Mark-Up Dynamics: Evidence from Spanish Firms". *Journal of International Economics Vol. 76*(1), pp. 107–115.

Shatz, H. J. (2003). "Gravity, Education, and Economic Development in Multinational Affiliate Location". *Journal of International Trade and Economic Development Vol. 12*(2), pp. 117–150.

Silvis, H. and R. Lapperre (2010). "Market Price and Quota: Half a Century of CAP Experience". In A. Oskam, G. Meester, and H. Silvis (Eds.), *EU Policy, for Agriculture, Food and Rural Areas*, Chapter 8, pp. 165–182. Wageningen: Wageningen Academic Publishers.

Social-Security. *Muestra Continuas de Vidas Laborales (MCVL).* (Accessed on April 2015) Ministerio de Empleo y Seguridad Social, Madrid.

Social-Security (2004, 2012). *Autónomos, Propiamente Dicho, en Alta en la Seguridad Social.* Ministerio de Empleo y Seguridad Social, Madrid. Available online at: http://www.empleo.gob.es/es/sec_trabajo/autonomos/economia-soc/autonomos/estadistica/.

Social-Security (2015). *Muestra Continuas de Vidas Laborales. Guía del Contenido.* (October 10, 2015) Ministerio de Empleo y Seguridad Social, Madrid.

Soloaga, I. and L. A. Winters (2001). "Regionalism in the Nineties: What Effect on Trade?". *The North American Journal of Economics and Finance Vol. 12*(1), pp. 1–29.

Staab, A. (2013). *The European Union Explained.* (3rd ed.). Bloomington: Indiana University Press.

StataCorp (2011). *Stata Statistical Software: Release 12.* College Station, TX: StataCorp LP.

Stead, D. R. (2007). "Common Agricultural Policy". *EH.Net Encyclopedia*. (June 21, 2007) Edited by R. Whaples. Available online at: https://eh.net/encyclopedia/common-agricultural-policy/.

Tinbergen, J. (1962). *Shaping the World Economy.* New York: The Twentieth Century Fund.

UN (2006). *Standard International Trade Classification. Revision 4.* Statistical Papers, Series M No. 34/Rev. 4. Department of Economics and Social Affairs, United Nations, Geneva. Available online at: http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=28.

UNCTAD (2014). *Bilateral FDI Statistics.* (Accessed on October 2015) Division of Investment and Enterprise (DIAD), United Nations Conference on Trade and Development, Geneva. Available online at: http://unctad.org/en/Pages/Home.aspx [Investment and Enterprise / FDI Statistics].

Westerlund, J. and F. Wilhelmsson (2011). "Estimating the Gravity Model without Gravity using Panel Data". *Applied Economics Vol. 43*(6), pp. 641–649.

WineAccess. "Tempranillo". (Accessed on January 2015) Ardmore, PA: WineAccess, Inc. Available online at: http://www.wineaccess.com/wine/grape/tempranillo.

Wooldridge, J. M. (2006). *Introductory Econometrics* (3rd ed.). Mason, OH: Thomson South-Western.

Zobbe, H. (2001). *The Economic and Historical Foundation of the Common Agricultural Policy in Europe.* Paper presented at Long Term Changes in the European Economies. Conference of the European Historical Economics Society, Oxford.