Università degli Studi di Cagliari

DOTTORATO DI RICERCA

SCIENZE ECONOMICHE ED AZIENDALI

Ciclo XXIX

TITOLO TESI

Trade costs, a twofold empirical analysis

The persistent effect of Roman roads on Italian provincial trade costs and the determinants of country trade costs by geographies

Settore scientifico disciplinare di afferenza

SECS-P/01

Presentata da	Vania Manuela Licio
Coordinatore dottorato	Prof. Andrea Melis
Tutor	Prof. Rinaldo Brau
Cotutor	Prof.ssa Anna Maria Pinna

Esame finale anno accademico 2015 – 2016

Tesi discussa nella sessione d’esame marzo – aprile 2017

Trade costs, a twofold empirical analysis

The persistent effect of Roman roads on Italian provincial trade costs and the determinants of country trade costs by geographies

Vania Manuela Licio

Abstract

Starting from early two thousands, the interest of trade literature on trade costs has grown exponentially. Theoretical and empirical studies have focused on the importance of trade costs as main drivers in international economics and as one of the major causes that impedes trade between and across nations.

Trade costs matter. Decline and rise of trade costs over time describe trade booms and busts in the past 150 years. They play a fundamental role in economic welfare, growth and development. And the research on the structure of trade costs, on the best method to measure them and on what determines trade costs is far from being complete.

This thesis is completely devoted to the trade costs subject and, in particular, to the indirect method of computing trade costs according to the 'top-down' approach proposed by the pioneering work of Novy in 2013. In order to provide the best overview on this theme and in order to add new knowledge to the trade costs literature, the research behind this thesis has been structured in three main chapters: one theoretical and two empirical. One of the two empirical applications includes also an instrumental chapter which gives account of a new data set constructed ad hoc for the first empirical investigation.

The first chapter is the theoretical support of the following two empirical analyses. It presents a detailed survey on trade costs from different perspectives, focusing on the indirect measure of trade costs, the dependent variable exploited in both following empirical investigations. The chapter leaves space for new thoughts on the indirect measure of trade costs, adding new reflections to the existing literature.

The first empirical application looks at the persistent effect of the Roman road network on current Italian provincial trade costs. This investigation presents three novelties. First, it is the earliest that studies the long-term effects of Roman roads using a measure in kilometres by province created specifically for this empirical application. Second, the approach exploited to measure Roman roads at the Italian provincial level can be easily extended to other more detailed or simpler investigations, at country or different sub-country levels and to one or more countries. Since the importance of this new data set, an entire chapter has been devoted to its presentation and explanation. Third, and most important, the empirical application on the persistent effect of the Roman road infrastructure is the first, to the author's knowledge, that uses an indirect measure of trade costs at the provincial level. The main idea behind the chapter is that history matters and that the Roman road infrastructure is affected by a legacy that lasted for more than two thousand years.

The second empirical application investigates on the best sources of trade costs, using a large data set of possible trade costs determinants, exploring these determinants by geographies of countries and for a wide number of nations. The novelty of this analysis lies exactly in the topic of the chapter. Due data constraints and measurement problems, the investigation on the determinants of trade costs is hard. It is harder if considering a large sample of countries and it is harder if distinguishing countries by geography. In this chapter, geography is measured in terms of 'degree of insularity'. The main idea behind it, is that geography matters. It matters when measuring trade costs and it matters when assessing the main sources of trade costs. The explanation of the main sources of trade costs by geographies of countries allows to better address policy makers in defining strategies and solutions to make trade less costly.

ACKNOWLEDGMENTS

First and foremost I would like to express my gratitude to my advisor Prof. Anna Maria Pinna for the continuous support, for her patience, motivation and enthusiasm during our meetings and calls. Her supervision helped me during all the time of my research and she has been a brilliant mentor for my Ph.D research. This thesis would not have been possible without her precious guidance. I am also in debt with Prof. Rinaldo Brau for his constant presence during these three years and for his invaluable commitment as tutor.

My sincere thanks go to Prof. Luca De Benedictis and Prof. Nicola Coniglio for the time they spent in reading this thesis, for the accurate and detailed reviews they provided, for the helpful comments, suggestions and ideas they gave to improve the current work also with future research.

I would like to thank Prof. Luca De Benedictis for sharing with Anna and me his first idea on the persistent effects of Roman roads on modern economic outcomes and for having given me the opportunity to freely develop this theme. It has been a pleasure to work on this fascinating and motivating idea and to devote a part of my Ph.D thesis to it.

This thesis has also benefited from the advices of Prof. Italo Meloni, Prof. Emanuela Marrocu and Prof. Fabio Cerina. I thank Prof. Emanuela Marrocu and Prof. Italo Meloni for the positive and supporting environment for research they guarantee in CRENoS and CIREM.

I thank the participants at the VII Workshop on Institutions, Individual Behaviour and Economic Outcomes in Alghero (June 2016), at the 7th International Conference Economics of Global Interactions: New Perspectives on Trade, Factor Mobility and Development in Bari (September 2016) and at the 57th Annual Conference of the Italian Economic Association in Milan (October 2016) for the helpful comments and discussions provided to the work "The long-term effects of the historical Roman road network: Trade costs of Italian provinces".

I would like to thank all my colleagues in CRENoS for the time spent together and for the stimulating discussions.

Last but not the least, I would like to thank my family, my parents and my sister Sara, for all their love, faithful support and spiritual encouragement during my Ph.D study, throughout writing this thesis and all my life in general. Thank you.

Vania Manuela Licio gratefully acknowledges Sardinia Regional Government for the financial support of her PhD scholarship (P.O.R. Sardegna F.S.E. Operational Programme of the Autonomous Region of Sardinia, European Social Fund 2007-2013 - Axis IV Human Resources, Objective l.3, Line of Activity l.3.1.).

CONTENTS

INTRODUCTION

Trade costs are influenced by geography and history.

While the economic literature has highly underlined how historical facts and geographical features persistently affect current economic outcomes today as in the past, the research has not extensively focused on what are the effects on trade costs.

Trade costs matter. They matter because they determine the pattern of bilateral trade and investment, as well as the geographical distribution of production, and because they determine countries' ability to take part in global trade networks. Countries with high trade costs are likely to have more expensive imports and less competitive exports, making their participation in the global international supply chain difficult and missing the resulting growth in trade and incomes. Although in the last three decades trade costs have extremely decreased driving a consequently increase in international trade flows, trade costs remain high, have welfare implications and play a big role in the political agenda of nations.

The common wisdom suggests that history and geography should be taken as given, since they are not affected by current aspects or they are not human-determined. But this thought is not completely correct. The deep influence of the government and the big impact of political actions on trade costs make historical and geographical factors active elements in the trade costs field.

On this framework, the core idea of this thesis is to present a study on trade costs providing a review of the most influential literature and a specific twofold empirical analysis with a dual spirit: one entirely historical and the other partially geographical.

The first chapter provides a survey on trade costs, focusing on the 'top-down' approach proposed by Novy (2013). It represents the common theoretical support for the next two empirical chapters, which employ an indirect measure of trade costs as the key economic outcome variable of the analysis. The idea is to provide, on the one hand, a coherent, general and complete overview of the whole trade costs subject, producing, paragraph by paragraph, a rational discussion on the main insights of the trade costs issue that will serve as a theoretical base for subsequent empirical applications. On the other hand, it presents some 'new thoughts' on trade costs. This paragraph, that represents the originality of the chapter, discusses in a constructive critical way the measure that will be exploited in both empirical analyses, leaving space for future research on these thoughts.

Chapter 1 underlines how, starting from the first era of Globalisation, the literature has retrieved in trade costs the key reasoning behind trade booms and busts of last 150 years. One of the main divergences in international trade lies in the form of trade costs. For more than fifty years, trade costs have been modelled in a multiplicative form, implying that trade costs were proportional to the value of the traded good. But, new empirics suggest that the pattern of trade costs is more resembling to a per unit cost. Despite their importance as drivers of the geographical pattern of economic activity around the world, research on trade costs remains limited. Nevertheless, there have been many attempts to understand what trade costs include, how important the single elements are, and how to measure trade costs. Much effort has focused on the direct measurement of various trade cost components, such as international transport costs. By summing all the elements together, the 'bottom-up' approach can produce an estimate of the overall level of trade costs faced by exporters and importers. However, this approach does not provide an all-inclusive measure of international trade costs, and combining the different measures and indicators into a comprehensive measure is hardly feasible because of data availability and unobservable measures. More recently, another strand of research, relying on the structural gravity model, has attempted to obtain a comprehensive index, inferring it from national productions and trade between countries. This indirect approach, aimed at producing a 'top-down' estimate of trade costs, tries to deduce implied trade costs from trade data, without specifying a trade cost function, and is the key approach used to measure the dependent variable of both empirical applications of this thesis.

The second and third chapter focus on history and historical infrastructures. They take into account a particular and significant historical feature and period: the Roman road network and the greatness of the Roman domination at the peak of its empire.

Chapter 2 represents the novelty of this thesis, since it presents the new measure of Roman roads that has been constructed at the Italian provincial level: it computes the length in kilometres of Roman roads for each province in Italy. This measure contributes to the literature on historical infrastructures, providing a new precise measure to use for empirical purposes, easy to extend at the regional or at the country level and simple to replicate in all those territories where Roman roads have been constructed. In addition, it demonstrates the relatively recent interest of economic research on the long-term effects of history on present economic outcomes, providing an extended review of the main contributions in literature and presenting a focus on four more specific branches: the persistent effects of Italian history, the persistent effects of historical infrastructure, the persistent effects of the Roman road network and the case of historical episodes that have not had long-term effects.

Chapter 3 contributes to the literature on the heritage of history, providing an empirical application of the new created Roman road measure. The aim of the chapter is to explore the connection between the indirect measure of trade costs and the historical Roman road network for the sample of 107 Italian NUTS3 provinces, testing whether differences on contemporary trade costs between provinces can be traced back to the long lasting impact of the Roman road system. Italy represents a perfect study case for two reasons. First, Italian contemporary territory was completely under the Roman empire and almost all provinces (108 out of 110) include Roman roads within their area. Second, Italy is characterised by a lasting duality between the economically developed North-Centre and the less developed South. On the one hand, the extension of the Roman road network on the entire peninsula allows to evaluate the Italian economic duality as not originating from a different Roman road equipment: the southern provinces had a denser Roman road system. On the other hand, the total dominion and supremacy of the Romans in the Italian territory does not enable to perform a counterfactual analysis at the Italian level. To implement a similar control, the observation background should be changed and extended: future research will be addressed to this direction. A dual inferential analysis is adopted. First, the simple and comprehensive effect of the Roman infrastructure on a current measure of trade costs is assessed. Second, Roman roads are used as an instrumental variable for the current infrastructure to investigate whether old infrastructures left a mark on present infrastructures. Results suggest that, having an integrated system of roads, as it was during the Roman domination, plays an important role on the current trade. Provinces with a large Roman road network are more prone to having less trade costs and, therefore, tend to trade more abroad than with themselves, according to the 'top-down' approach proposed by Novy. On the whole, the study confirms not only the importance of history in the contemporary economic development, through past better institution, but also the main role of history in shaping a more open mentality of peoples.

In the fourth chapter the focus is on geography and on what makes trade costly. Exploiting the classical bilateral indirect measure of trade costs at the country level, the aim is to provide an empirical investigation on what are the determinants of trade costs by different types of geographies of countries.

Chapter 4 contributes to the literature on trade costs and on the importance of geography for economic performance and development. Research on the determinants of trade costs has been very little investigated, due data constraints that typically affect this type of study. Nevertheless, the question is relevant, since it helps policy makers to take the enhanced strategies to reduce trade costs between countries. Empirical investigations on whether the determinants differ across countries' geography might produce more accurate policy recommendations: it has been highly underlined how geography may represent a barrier to trade and, in general, a disadvantage in economic terms. On these bases, trade costs are computed for 188 countries, distinguished in four geographical (insular) categories according to Pinna and Licio's (2013) insularity data set: landlocked countries, coastal countries, partial-insularity countries and island-states. The potential main determinants of trade costs are, instead, selected within four key spheres and include: geography, logistics, competitiveness and connectivity, infrastructure; documentary, border and transport compliances. The use of the Heckman approach together with the ‘unadjusted’ indirect measure of trade costs in the main empirical analysis, allows to explore two aspects of the trade costs subject: on the one hand, the participation to trade and, on the other hand, the extent of trade costs. Results suggest that logistic abilities and competitiveness qualities are the main determinants of trade costs for all geographical groups. The shipping connectivity has a more marginal role and matters more for coastal countries than for other categories. Its role is not so essential for the island-states group. Within the geographical sphere, it is mainly the physical distance between partners that is important. However, whereas it is significant in determining the extent of trade costs, its role in influencing the involvement in global exchanges is secondary. Surprisingly, the quality of access to seaports is important for the landlocked countries. Within the island-states, instead, it is more the quality of the air infrastructure that matters. Border compliances and red tapes marginally affect trade costs and, compared to other factors, they are irrelevant. However, for island-states group the effect of cost per container is particularly important in determining the size of trade costs. Data availability represented the main difficulty of the analysis, since it plagued the variability within some geographical groups as well as the inference analysis. Further research is needed in order to have more precise estimates and to capture the phenomenon in a more precise manner. In light of this, since problems of model uncertainty related to the best specification and the selection of the explanatory variables, future analysis will exploit Bayesian Model Averaging (BMA).

Chapter 1

A SURVEY ON TRADE COSTS

1.1 Introduction

During last fifteen years, a growing number of theoretical and empirical studies started to focus on the subject of trade costs, documenting their importance as one of the main drivers in international economics and as one of the major causes that impede trade between and across nations. Trade costs affect the economic development and are the key factor describing why some countries grow and diversify and why other ones don't. Even if growth and development of countries cannot be explained just relying on the trade costs reasoning, trade costs are fundamental in determining nations and consumers' welfare.

The reduction of nations and consumers' welfare, linked to high trade costs, follows the conventional wisdom. Anderson and van Wincoop (2002) suggest that welfare implications, because of policy-related trade costs, are large. The World Trade Organisation (WTO) calculates that high trade costs make consumer prices higher and place uncompetitive producers out of the global market, reducing the resulting economic welfare. Lower trade costs are linked to lower net poverty, even though the extent of the reduction varies across countries. Moreover, lower trade costs lead to improved production efficiency and to enhanced performance. On the contrary, higher trade costs impede countries to take advantage of specialisation by comparative advantage, isolate them from world markets, and hamper the access to technology and to intermediate inputs, limiting the benefits of being involved in global value chains. Trade costs are not able alone to explain why some countries are developed and why other countries lie in a poor economic condition, but, combining the information coming from trade costs with the information deriving from other factors, it is possible to understand the difficulties faced by some countries in following a sustainable growth and in taking advantage from their comparative advantages.[1]

According to Anderson and van Wincoop (2004, p. 691) "*Trade costs [...] include all costs incurred in getting a good to a final user other than the marginal cost of producing the good itself: transportation costs (both freight costs and time costs), policy barriers (tariffs and nontariff barriers), information costs, contract enforcement costs, costs associated with the use of different currencies, legal and regulatory costs, and local distribution costs (wholesale and retail)*". Basically, all those costs linked with the trade of goods across national borders and which hamper trade between countries are trade costs. In this view, trade costs can be seen as a measure of distance between markets and as the major obstacle to international economic integration. Nevertheless, as highlighted by Novy (2013), trade barriers matter not only when considering international trade across countries, but for domestic trade as well. Following Chapter 3 explores a particular form of trade costs, measuring trade costs at the Italian provincial level and examining whether higher trade costs represent for Italian provinces a constraint for trading with foreign countries.

International trade has experienced one of the fastest growths over the last few decades in terms of theoretical discovers. The importance of international trade for economic growth has been pointed out since Adam Smith (1776) and has driven precise research. Indeed, during the Sixties, the opening of international trade and the need to account for sources of growth, besides the traditional ones, led to a growing research on trade and economic growth.[2] The further increase of international exchanges and the changing character of trade during the subsequent years provides an extra incentive to deepen research. This growth in research is strictly connected to the increase that international trade experienced in terms of flows;[3] and the empirical research has proved that one of the sources of increasing trade must be found in the decline of international trade costs.[4]

Starting from the Eighties, countries started to become more interdependent, interconnected and integrated, and trade started to change as it never did during the previous centuries. The changing character of trade and the Globalisation have been fundamental during this process.

Globalisation and the nature of trade are deeply correlated, but not coincident in terms of periods. Until 1850 trade was characterised by a high specialisation and mostly by low

[1] OECD/WTO (2015).
[2] Afonso (2001).
[3] In 1970 total exports were 12 percent of GDP, in 2006 they amounted to 31 percent (Johnson, 2008).
[4] Novy (2013).

volumes of trade in goods with high value per weight (spices, precious metals, etc.). The high specialisation lasted for more than a century, from 1850 to 1980. Britain first, then Europe and the U.S.,[5] and later the OECD countries represented the guidance of the world with growing levels of trade in bulk commodities and manufactures. It is from the 1980 that the world witnesses a completely radical change: de-specialisation, trade in intermediate goods, intra-industry trade and trade in services are the four traits of the new character of trade that lasts until today. On the other hand, the globalisation eras have been largely studied to understand what caused trade booms during the last 150 years and what in particular determined the increase in trade during the last three decades. Although there is not a common agreement about the number, the beginning and the end of each era of Globalisation,[6] literature strongly confirms that during the globalisation process countries and firms become more integrated. They achieve the full economic integration when the domestic economic elements are able to affect home and foreign market in the same way. In that sense, Fouquin and Hugot (2014) suggest that, in terms of trade costs, countries reach a perfect economic integration when bilateral trade costs are equal to domestic trade costs and the measurement of relative trade costs across years provides a satisfactory assessment of the extent of trade globalisation.

While every globalisation era had its personal characters in terms of what drove the boom, the research has proved that the fall of trade costs seems to be the common factor along the entire globalisation process.

Late nineteenth-early twentieth century Globalisation (1870-1913) found in the new technology (steam and internal combustion engines, telegraph, electricity) and in the liberal trade policy its two main drivers. Great Britain was the leading economy in the world. These were the years of the European Colonialism and of the big migration of Europeans to the New World. Capital was also free to move between countries: financial integration was higher than today. Trade in services was still limited and goods were mainly moved by ship.

---

[5] Baldwin and Martin (1999) highlight how the increased openness of the U.S. in terms of foreign exchange transactions after the World War II is one of the distinctive marks of the second wave of Globalisation.

[6] Baldwin and Martin (1999) consider two waves of Globalisation (1820-1914 and 1960-present). De Benedictis and Helg (2002) identify three periods (end of nineteenth century-1914, 1945-1980, end of twentieth century). Jacks et al. (2008) look at two globalisation waves: the first era from 1870 to 1913, the second dating from 1950. Johnson (2008) classifies the globalisation process into three eras: 1860-1914, 1944-1971 and 1989-present. Fouquin and Hugot (2016) identify two waves of Globalisation: the First Globalisation of the late nineteenth - early twentieth century and the post-World War II Second Globalisation, underlying that the precise timing of the first era remains unclear. In a prior work (2014), Fouquin and Hugot assert that the nineteenth century trade globalisation began roughly twenty years before the 1860s, in conjunction with the trade liberalisation.

The debate in literature asserts that the transportation improvement and the consequent fall of transportation costs was decisive in determining trade globalisation prior to 1913. The new technology of steam engine and the adoption of steamship promoted trade after 1870. While steamships were responsible for spurring trade cross-country, railroads spurred trade within-country. On this view, the contributions of Pascali (2014), Frieden (2007), James (2001) and Harley (1988) are the most influential in supporting the idea that the first era of trade globalisation was due to the decline in transport costs. Other literature, Estevadeordal et al. (2003) and López-Córdova and Meissner (2003), is more prone to explain the increasing trade in the late nineteenth-early twentieth century using the monetary regime reasoning. Estevadeordal et al. (2003) confirm that the fall in transport costs was one of the main forces behind the First Globalisation, but they suggest how the rise of the gold standard had an important role in the increasing trade. On the same view, López-Córdova and Meissner (2003) find that the gold standard regime had a positive effect on trade flows: two countries on the gold standard would trade 60 percent more with each other than they would do with countries belonging to other monetary regime. Jacks (2006), looking at the North Atlantic grain markets between 1800 and 1913, finds that changes in freight costs can explain just a small part of the variation in trade costs in those markets, suggesting that the monetary regime, the commercial policy and the political environment strongly influence trade costs. From a different perspective, O'Rourke and Williamson (1999) find that the increase of commodity trade through the Atlantic Ocean in the late nineteenth century was due much more to the decline in transport costs than to the liberal trade policy, but underline how the increase of income and the import demand had a main role in fostering trade. On the same line, Estervadeordal et al. (2003) and Jacks et al. (2011) highlight that the growing trade flows were mainly determined by the increasing GDP (Gross Domestic Product) and by the combination of lower transportation costs and the gold standard system. Conversely, Fouquin and Hugot (2016) show that the decline of international trade costs started in the 1840s, contradicting the studies that ascribe the First Globalisation to new technologies or to the gold standard. An earlier work of Jacks et al. (2008), using data on GDP and total trade flows for major economies, finds that the decline and rise of trade costs over time describe trade booms and busts in the past 150 years. 55 percent of the pre-World War I trade boom is explained by the decrease of trade costs, 33 percent when considering the post-World War II; similarly, the steep increase in trade costs during the Great Depression explains the interwar trade bust. Jacks et al. (2011) show, in fact, how in the period 1870-1913 the decline in trade costs was the main factor explaining the increasing global trade. In their work they decompose the average growth[7] of international bilateral trade for 130 countries. The contribution of the decrease of trade costs accounted for 290 percent of the total 486 percent average growth of international trade (about 60 percent when the average growth is expressed as 100 percent); the growth in output, the growth in income similarity and the change in multilateral factors contributed respectively for 225 percent, -11 percent and -18 percent. After the World -War II (from 1950 to 2000), the increasing in trade was 484 percent and the decline of trade costs accounted for 148 percent (about 30 percent when the average growth is expressed as 100 percent).

The second era of Globalisation corresponds to the period starting with the Bretton Woods Agreement and concluding with the end of the gold standard (1944-1971). Although new technological progresses (television, communication via satellite, jet planes and container), the main drivers of the economic integration were represented by the international regulations and organisations: the World Bank (WB), the International Monetary Fund (IMF) and the General Agreement on Tariffs and Trade (GATT) were all founded during this period. The U.S. became the leading economy in the world and the dollar the identifying monetary basis of the financial system: the Bretton Woods Agreement established for each national currency the exchange with the U.S. dollar, that was in turn fixed to the gold standard. The movement of goods, capital and people were regulated even if trade tariffs were gradually reduced thanks to the GATT. These were the years of the Cold War. With the Vietnam War and oil crisis in the early 1970s, the Bretton Woods Agreement got into a crisis and the gold standard system was abandoned.

The fall of the Berlin wall in 1989 marked the start of the current globalisation period (1989-present). The end of the Cold War and a more relaxed and peaceful political environment led to less trade barriers and to the constitution of the World Trade Organisation (WTO) in 1995. The Third Globalisation wave also witnessed the birth of personal computers, microprocessors and mobile telephones, which led to new trade and business. New economies emerged in the world scenario: China and Southeast Asia economies started to become important trade partners,[8] although Europe and the U.S. still represent the most influential and important actors in the global trade. Free trade and free movement of capital were the rule, but the free circulation of peoples was regulated. These are the years of the European Union constitution and enlargement and of the Schengen Agreement.

---

[7] The average growth is weighted by the sum of the two countries' GDPs in each pair to mitigate the influence of country pairs which trade rarely or irregularly (Jacks et al., 2011).

[8] These countries are named by the World Bank (2002) '*Globalizers*'.

The importance of trade costs lies also in economic theory. Trade costs seem to be the plausible common key solving the six major empirical puzzles in international macroeconomics. Obstfeld and Rogoff (2000) address these puzzles by introducing iceberg trade costs into the most benchmark models of international macroeconomics.

When considering the central question in positive trade theory '*Why do countries trade?*', the Ricardian principle of comparative advantage is strongly linked to the issue of trade costs. Dornbush, Fisher and Samuelson (1977) describe that impediments to trade, like tariffs and transports costs, determine which goods are not traded due to higher price levels. Assuming transport costs à la Samuelson (1954) and given country's profitability to produce commodities when the domestic unit labour costs are lower than the foreign unit labour costs adjusted for the iceberg transport costs,[9] Dornbusch and co-authors are able to distinguish the range of non-traded goods from the one of traded goods. If the relative wage increases, the fraction of goods produced domestically decreases and the range of imported commodities increases. In other terms, they show how transfers affect the equilibrium relative price and, in turn, the share of goods that are traded and not. Accordingly, Yaylaci (2013) remarks that, with high trade costs, goods are more costly in foreign countries, exports become uncompetitive and the relative comparative advantage is cancelled out. The comparative advantage is also affected by trade in intermediate goods. Even if two countries have same trade costs, using intermediate goods for which trade costs in the intermediate goods market are lower for a country than for other countries (for instance because the first country is closer to a low-priced source of some intermediate goods), offers a gain in terms of comparative advantage.[10] The third effect of high trade costs lies in the geography underlying the comparative advantage. When trade costs are absent, countries' exports are determined only by their relative comparative advantage and their imports only by their preferences. Differently, with (high) trade costs, countries would trade more internally than internationally, preferring closer or neighbouring partners than faraway countries. More in deep, the interaction between comparative advantage and trade costs is found to determine the degree of specialization of a country (Markusen and Venables, 2007): economies with high trade costs cannot benefit from specialisation and trade, with consequences on economic development. Pomfret and Sourdin (2010) remark how high trade costs generate a disadvantage in trade and income growth due the exclusion from international supply chains. Hoekman and Nicita

---

[9] Dornbusch, Fischer and Samuelson (1977) rely on a two-country Ricardian model with a continuum of goods.

[10] "*High trade costs deny firms access to technology and intermediate inputs, preventing their entry into, or movement up, global value chains*" ("Aid for Trade at a Glance 2015: Reducing Trade Costs for Inclusive, Sustainable Growth, OECD", WTO 2015, p. 35).

(2011) show that non-tariff measures and domestic trade costs represent relevant barriers to international trade. Rubin and Tal (2008) argue that transportation costs represent a bigger impediment to trade than barriers originating from policy decisions, such as tariffs.

The importance of trade costs lies exactly in the fundamental role they play in economic welfare, growth and development. Anderson and van Wincoop (2004) estimate that trade costs faced by a representative developed country are around 170 percent of the producer price of exported goods. For developing countries the percentage is even greater, since trade barriers are higher than those in industrialised countries. Poor countries face high absolute levels of trade costs compared to developed countries. As observed by Arvis et al. (2013a), they have higher tariff and nontariff barriers, poor and limited infrastructure systems, transportation networks and logistics service. Rich countries assist to faster decreasing trade costs trends; in developing states, although slow declining trade costs process in manufacturing, in agriculture the situation is static (Arvis et al., 2013a). The brief reasoning provided should give the sense of why trade costs matter and of why the trade costs issue should be taken seriously. To put into effect measures focused on reducing trade costs should be on the top of the political agenda of developing countries, in order to drive them away from marginalisation and include in the global trading economy. Trade costs are not the only reason behind underdevelopment and low welfare of countries, but, although their decline during last decades, they still matter and are high.

This chapter consists of eight paragraphs. In Paragraph 1.2, the recent debate on the structure of trade costs is examined to understand which form between the ad valorem and the specific structure better represents that of trade costs. Paragraph 1.3 compares the two main approaches for computing trade costs, direct and indirect, underlying strong points and disadvantages. The indirect approach proposed by Novy in 2013, which is used to compute the dependent variable used in the two empirical chapters, is exhaustively examined in Paragraph 1.4. Paragraph 1.5 presents some (new) thoughts on the indirect measure of trade costs. A brief evaluation of the elasticity of substitution, important element in the trade cost equation, is analysed in Paragraph 1.6. Paragraph 1.7 provides an investigation on the main and the most influential determinants of trade costs, in order to examine the sub-components of trade costs and their relevance. In Paragraph 1.8 some concluding remarks are provided.

1.2 Multiplicative or additive trade costs?

Recent research has grown some criticism about the structure of trade costs. As pointed out by Irarrazabal et al. (2015), since Samuelson (1954), literature in international trade has modelled and perceived trade costs in *multiplicative* terms, deducing that trade costs are proportional to the monetary value of the goods traded and implying that goods with higher prices are more expensive to trade. This conjecture was mainly due to the methodical convenience of multiplicative form. New empirical research departs from this view, suggesting that trade costs sketched in *additive* terms also have a role in explaining the character of trade costs and the pattern of trade, and that an important part of the transportation cost is additive rather than multiplicative.

Multiplicative trade costs (also defined *iceberg* or *ad valorem*) are expressed in the form of a percentage of the traded good without shipping costs (i.e. the free on board costs, f.o.b.[11]). The specification assumes that, to deliver one unit of a good to the foreign destination, the exporting country should send *t* units of the good, because *t-1* units are 'melted' during the transportation. The basic idea behind the original formulation of Samuelson (1954) was that trade flows are reduced by frictions, and he used the comparison with icebergs to clarify his point: as icebergs melt in the ocean, traded goods melt their value during transportation.

Additive trade costs (also named *specific* or *per unit*) are structured as a constant cost applied to every unit of the traded good.

Trade costs are strictly linked to the patterns of trade, to comparative advantage and to welfare, and the way according to which they are modelled are related to these issues as well. Irrarazabal et al (2015) find that reductions in specific trade costs yield larger welfare gains than reducing trade costs in the ad valorem form.[12] Differently, Cole (2011) suggests that higher welfare gains are achievable when reducing ad valorem trade costs rather than per unit ones.[13]

The form of trade costs adopted determines how well single trade impediments are captured, is able to explain why some firms find not profitable to enter foreign markets and influences what goods are traded. The discussion on the shape of trade costs becomes particularly

---

[11] The f.o.b. acronym stands for free on board and is mainly intended for sea transport. A free on board cost includes all charges up to move the goods on board a ship in a determined point of origin (usually a port). These expenses are borne by the exporter (seller). The remaining costs for delivering the goods to the final destination are paid by the importer (buyer).

[12] Irarrazabal et al (2015) exploit a model that allows for heterogeneity across marginal costs.

[13] Cole (2011) adopts a monopolistic competition model with heterogeneous firms and homogeneous marginal costs.

significant in terms of heterogeneous goods, acquiring different useful insights for the trade costs issue.

In relation to non-differentiated goods, Deardorff (2014) develops a simple example where, supposing a world with four countries and a homogenous good, he shows how trade costs play a big role for local comparative advantage. In order to demonstrate his argument, he exploits per unit trade costs for the good to get it from a country to another. The use of iceberg trade costs, although possible, is not suitable since, as highlighted by Deardorff, the value of the ad valorem trade costs changes with the price of the traded good. With his example Deardorff shows that, if trade costs are larger than the difference between autarky prices of two adjacent economies,[14] then no trade will be possible between those countries. But, if trade costs are very low, so that the difference between autarky prices of two economies is larger than trade costs, then the country with the smaller autarky price will export to the other one. When trade costs become smaller and smaller and go to zero, the world reaches a unique-equilibrium price and all four countries are involved in trade: countries with higher relative autarky prices import, countries with lower relative autarky prices export.

However, the form of trade costs has important implications when considering differentiated commodities and different-quality goods. During last years the Alchian-Allen theorem, developed more than fifty years ago, has reacquired interest in international trade literature. Recent research has provided new evidence on the Alchian-Allen effect and on the shape of trade costs.

Specific trade costs reduce the relative price of high-quality commodities. In 1964 Alchian and Allen pointed out that, due the transportation costs,[15] firms are tempted to trade high quality products in foreign markets, while domestic trade and demand are centred on low-quality goods. This idea is at the core of the Alchian-Allen effect theorem, one of the most classical theorems in international trade. The theorem is also known as the '*shipping the good apples out*' and asserts that, when the prices of two substitute goods, one of high quality and the other of low quality, are both increased by a fixed per unit freight charge, this cost will affect the high quality-good, lowering its relative price and raising the relative demand for it (Alchian-Allen effect). In other terms, the theorem implies that shipping costs of the iceberg form don't affect relative demands for quality, while per unit costs do. Although its

[14] Countries (named A, B, C and D) are located along a straight line at equals intervals (Deardorff, 2014).

[15] Transportation costs represent one of the most predominant elements of trade costs. See paragraph 1.3 for a discussion.

importance, interest of literature remained only on the theory and on the concept of the theorem, without providing an evidence in an empirical setting. Hummels and Skiba were the first ones, in 2004, to test the Alchian-Allen theorem by relating average prices, borne by the exporter for a specific good, to freight and tariffs. The originality of the paper is represented by the form in which Hummels and Skiba express the shipping costs: per unit rather than ad valorem. They underline how the classical version of the theorem implies that, when the transportation cost is expressed in the additive form, the higher the per unit freight is, the higher the demand for the high-quality products is. Ad valorem tariffs have an opposite effect: they reduce the demand for high-quality goods. In this view, the exporters select trading destinations changing the f.o.b. price of traded goods and importers can pay different prices for the same good. Starting from this fact, Hummels and Skiba (2004) aim first to test the Alchien-Allen theorem with trade data, secondly to study the relationship between additive and multiplicative trade costs, and finally to prove that the existence of per unit costs is not described by monopoly pricing-to-market attitude. Their estimation analysis is divided in two parts; for both sections they use bilateral trade data from 6,000 country pairs at the 6-digit level of the Harmonized System (HS), comprising information about prices, quantities, shipping costs, and ad valorem tariffs. The first part of the analysis estimates a transportation cost function, showing that these freight costs are more likely to be on the additive rather than on the multiplicative form. Moreover, the first section demonstrates the traditional formulation of the Alchian-Allen theorem stated above: transportation costs increase the demand for high-quality goods, tariffs reduce it. In fact, a doubling of shipping costs produces a consequent 80-140 percent raise in the f.o.b. prices. When doubling tariffs, the f.o.b. prices reduce by 140-250 percent. The second part of the analysis is aimed at understanding whether prices vary across destination due to quality or due to pricing-to-market.[16] Literature emphasises that price variation across markets comes from pricing-to-market, since monopolists are prone to change the price according with variations in tariffs or exchange rate. Hummels and Skiba (2004) demonstrate that pricing-to-market alone is not able to explain the extent and the size of price elasticity with respect to tariffs or to transportation costs. They show that trade costs are more of the additive than of the multiplicative form when goods have the same quality. Their estimates show that an increase in the per unit cost leads to an increase of the demand for high-quality goods, and the larger the price difference is, the more this effect exists. Differently, a raise in the ad valorem trade costs reduces the

[16] Pricing-to-market refers to the producer's inclination to charge different prices for the same product in the foreign and domestic market due international relative costs.

relative demand of higher-priced goods. Moreover, they explain the Alchian-Allen effect in terms of the variation in demand for quality determined by a change in the non-price portion (like distance, quantity shipped, etc.) of the freight charge. The sign and extent of the Alchian-Allen effect depend on the elasticity of the freight costs with respect to prices: when the elasticity is equal to 1, transport costs are of the iceberg form (ad valorem); when the elasticity is smaller than 1, the Alchian-Allen effect exists and becomes stronger as the elasticity decreases; when the elasticity is equal to 0, freight rates are per unit; and when elasticity is greater than 1, a reverse[17] Alchian-Allen effect exists. Hummels and Skiba estimate an elasticity of tariffs with respect to price of 0.6, providing evidence for the Alchian-Allen theorem and ruling out the iceberg assumption. On the whole, the work of Hummels and Skiba (2004) proves a significant empirical evidence for the Alchian-Allen theory and suggests that, differently from what international trade literature assumes, transportation costs are not of the iceberg form. However, they are not able to identify the magnitude of additive costs. The work of Irarrazabal et al. (2015), ten years later, tries to fill this gap. The novelties of their contribution are basically two. On the one hand, they look at the demand side, at how volume of exports changes when additive trade costs rise, given producer prices. This is aimed at the study of the size of per unit trade costs relative to ad valorem trade costs in a general equilibrium framework. The second contribution of the paper compares welfare consequences of ad valorem trade barriers with those of additive trade barriers. Irarrazabal et al. (2015) use Norwegian firm-level trade data (8 digit) of 21 foreign countries to estimate trade costs across firms for every product-destination pair in the sample (121 products and 21 foreign countries). They find that the elasticity of quantity demanded, given f.o.b. price when additive trade costs rise, is negative and increasing in the producer price (-0.04 for low price firms and zero for high price firms), providing evidence of the Alchian-Allen effect. In addition, they estimate that additive trade costs are 14 percent of the (median) f.o.b. price. When considering welfare implications of additive trade costs, they show that the welfare reduction due to an additive trade barrier is 2.02 percent; due to a multiplicative barrier, it is 1.31 percent. Basically, an additive import tariff reduces welfare around 50 percent more than an equal-tax-yielding multiplicative tariff.[18] The intuition behind this is that, additionally to reducing imported/home produced goods, it reduces relative imports of low/high priced varieties of goods. In other terms, additive trade costs affect

---

[17] "*[...] transport costs rise faster than goods prices so that an increase in the nonprice portion of the transport cost (e.g., distance) actually raises the price of high- relative to low-quality goods*" (Hummels and Skiba, 2004, p. 1391).

[18] An ad valorem tariff which yields the same total revenue of a per unit tariff (Irarrazabal et al., 2015).

relative demand both within and between countries, multiplicative trade costs only alter prices across markets; moreover, an additive tariff restricts trade more than an equal yielding multiplicative tariff. In summary, they conclude that empirics and theory should account for both iceberg and additive trade costs.

In other contributions, the form according to which trade costs are expressed is crucial in assessing other features of trade. In the work of Feenstra and Romalis (2014) the discussion between ad valorem and specific trade costs will serve to analyse the quality from the export unit values. The estimation of quality-adjusted prices in trade has become an important field in international trade literature, and recent papers suggest that quality allows to understand the patterns of trade between countries,[19] how countries specialise[20] and also provide useful information on the path of growth.[21] Feenstra and Romalis adopt a new methodology, combining non-homothetic preferences for quality and specific shipping costs. Their estimates suggest that richer countries export more qualitative goods than poorer countries. The quality-adjusted terms-of-trade declines, illustrating richer-countries' preference for higher quality goods. When considering exports, quality explains much of the variation in export unit values. From the imports perspective, instead, poor countries have lower quality-adjusted import prices. Therefore, export quality is more related to income than import quality. More on quality of internationally traded goods, Lugovskyy and Skiba (2015) aim to study a not previously researched issue: whether the level of quality specialisation is influenced by the geographical location of the country-exporter. To this end, they use a multi-country model based on the supply side behaviour of the producer to the demands from several destinations (importers), with '*arbitrary distributions of preferences for quality*' and with a functional form of the transportation cost, including both multiplicative and additive components. Trade costs are defined in three ways based on freight, insurance and distance. The results confirm how geography is important when considering the nature of exports in terms of quality: the proximity to richer trade partners increases exports quality and, according to the Alchian-Allen effect, exports to far-away destinations promote the production of higher quality products. Higher quality exports are positively affected by specific transportation costs and preference for quality, confirming the Alchian-Allen theorem also in multilateral trade. Moreover, the impact of bilateral trade costs is consistent with

[19] Hallak (2006).
[20] Schott (2004).
[21] Hummels and Klenow (2005).

theory: specific freight rates have a positive effect on the export prices, while for the ad valorem insurance rate the effect is negative.

A very recent work of Huberman et al. 2015 focuses on trade costs to assess what drove the Belgium's trade boom between 1870 and 1914. They test three hypotheses. First, the intensive margin[22] is affected only by ad valorem trade costs, whereas the extensive margin[23] is affected by both specific and ad valorem costs. Second, declines in trade costs increase bilateral trade as product differentiation increases and as firm-level heterogeneity decreases. Third, higher degrees of firm-level heterogeneity and product-differentiation mean larger impact of trade costs on productivity. To test the first two hypotheses they use a gravity model of bilateral exports employing Belgian exports data on manufactured goods between 1870-1910; data on distance, existence of trade agreements, etc., are used as proxies for specific and ad valorem costs. For the first hypothesis the dependent variable is represented by the number of products exported per country (extensive margin) and by the share of exports value relative to the number of products per country (intensive margin). The empirical analysis confirms that the intensive margin is influenced by ad valorem costs, while the extensive margin is determined by both specific and ad valorem trade costs. For the second hypothesis, the gravity model includes four different categories of goods based on increasing levels of product differentiation and separated by industry, to test effects of firm-level heterogeneity. Huberman et al. (2015) find that the impact of trade costs decreases as product differentiation and/or firm-level heterogeneity increase. To test the third hypothesis, they use also labour productivity and real output growth for 20 industries in two sub-periods (1880-1896 and 1896-1910) and the empirical analysis relies on OLS estimates of the relationship between export-growth and productivity growth. The results of the third hypothesis test suggest that industries with higher firm-level heterogeneity and product differentiation had a weaker relationship between export growth and productivity growth. The work of Huberman et al. (2015) includes some criticisms related to the data availability and reliability, the numerous assumptions in the quality estimation and for the general equilibrium analysis. In addition, in all gravity models they overestimate the impact of distance on trade and trade barriers.

It can be argued that that theoretical and empirical works should take both ad valorem and specific trade costs into account, since their different and specific features. On the one hand, additive trade costs increase demand for high quality goods, depend on distance and quality,

[22] The intensive margin refers to changes in trade values due changes in volumes of existing products to existing destinations.

[23] The extensive margin refers to changes in trade values due an increase of newly traded (or disappearing) products on diversification and/or new exporting destinations.

affect only the extensive margin and are considered in firm mark-up decisions. On the other hand, ad valorem trade costs decrease demand for high quality goods, depend solely on the f.o.b. and quality, affect both intensive and extensive margin and, like a tariff, increase the demand for home goods. When considering the Alchian-Allen theorem, the no iceberg assumption holds with an elasticity of freight rates with respect to price of 0.6. The general equilibrium analysis shows that the welfare decline due to an additive import tariff is 50 percent higher than in case of an equal revenue-yielding multiplicative tariff. Moreover, additive trade costs generate a larger decrease in trade flows. Richer countries produce and export more high-quality goods than poorer countries. In fact, as stated by Lugovskyy and Skiba (2015, p. 156), "*quality increases [...] in trade-weighted specific transportation cost to destinations other than the importer and [...] in trade-weighted preference for quality of destinations other than the importer*".

1.3 Direct vs indirect approach

Trade costs can be measured using two different approaches: direct and indirect.

The direct methodology collects and adds up direct available and observable data or proxy on individual trade cost factors to obtain a single measure. The main difficulty connected to this methodology lies in the availability, accuracy and incompleteness of data. Anderson and van Wincoop (2004, p. 692) stress that "*direct measures are remarkably sparse and inaccurate*". Accordingly, Novy (2013, p. 104) underlines that "*direct measures for appropriately averaged trade costs are generally not available*". The direct method is also called '*bottom-up*', because trade cost sub-components are added up to obtain the overall measure. Application of a 'bottom-up' approach necessarily requires the use of a trade cost function to relate the unobservable sub-component costs to the observable ones. The form of the trade costs function is strictly connected with the underlying economic theory. As pointed out by Anderson and van Wincoop (2004), the most straightforward form is multiplicative in its different components. Hummels (2001), instead, finds in the additive setup a more suitable structure.

The accessibility to accurate and complete sources of data is a fundamental issue, when talking about the direct measurement of trade costs, for two main reasons. On the one hand, to appreciate the contribution of observable factors, trade costs elements are typically used as explanatory variables in gravity equations of trade; on the other hand, when some information is not available, this causes an omitted variable bias problem, since the potential correlation between observable components included in the model with unobservable omitted trade costs. The limited data availability in the direct approach has been largely discussed in literature. Anderson and van Wincoop (2004) discuss the problem of the accessibility of data dividing trade costs in two main categories: policy-related (like tariffs, quotas, countries' trade policies, etc.) and environmental (transportation and time costs, insurance, wholesale and retail distribution costs, etc.) trade costs. They argue that the main difficulties linked to the policy barriers lie in the absence or fragmentation of data and in the aggregation bias which arises from the need, for some measures, to combine available with missing data. The coverage is scarce for some years or poor countries. Another problem comes from the way in which some sources report the trade-barrier information: some tariffs are expressed in specific rather in ad valorem terms. When considering nontariff barriers (NTBs), instead, Anderson and van Wincoop (2004) underline how problems of paucity and completeness arise, leading to the need of an indirect way to compute them. A typical approach consists in inferring NTBs from the comparison of prices or from trade flows. For trade costs linked to environment, the private nature of transport costs data seems to be the main limit in the computational analysis; however, some sources at the country level help for more detailed data. Time and transportation costs are strictly connected to goods and infrastructures, this makes them less comparable at a good-level analysis, but more comparable at the average level. The measurement of all other trade costs, like information barriers, contract enforcement costs or costs linked more to cultural or national aspects, is hard to achieve and, for some costs, not possible at all. Chen and Novy (2012) confirm how the limitedness of data represents the main obstacle for having a valid representativeness of trade costs and powerful direct measure. They divide trade cost sub-components in two main categories: explicit measures and implicit measures. The explicit measures consider and count the total amount of standards and regulations by country, industry, product and type of barrier. Chen and Novy (2011) clarify how different standards and technical regulations[24] represent a wedge between importers and exporters, since foreign producers are required to adapt their production to the importing country requirements, creating additional producing costs. In order to get rid of these costs, national governments try to harmonise standards and regulations between countries. The use of standards and regulations as trade costs measure is not straightforward. As noticed by Chen and Novy (2012) and other authors (like Swann, 2010 and Moenius, 2004), there are some drawbacks in using standards. Standards can be divided into national

[24] Technical regulations are also named Technical Barriers to Trade, TBTs (Chen and Novy, 2011).

and harmonised standards,[25] by type and by importance: simply adding up these standards, without considering such differences, can lead to inaccurate measures. Also regulations never represent a perfect measure. Chen and Novy (2012) observe how they can apply in one nation but not in the country partner, differently affecting trade across and between countries. Moreover, studies based on standards and regulations are not completely able to measure the real degree of barriers, since there is not perfect information about how rigorous and effective they are. On the other hand, the implicit measures calculate which are the effects on trade of standards and regulations, but they don't measure the total amount of them. Basically, implicit measures aim to establish how exports are affected by standards and regulations on exports, without assessing the size of such policies. On this view, the effectiveness and the stringency of these procedures are assessed through the competitiveness and the extent of exports.

Duval and Utoktham (2011a) about the direct measurement, stress how the literature during last years has put much effort on trying to calculate direct trade costs. Some authors have basically focused on international transport costs. Limão and Venables (2001) measure shipping costs of firms for a standard container from Baltimore to different countries and use also the c.i.f. [26]/f.o.b. ratio as a measure of transport costs. On this point, Hummels (2001) underlines how researchers must be aware that the comparison of the free-on-board (f.o.b.) data, reported by the exporting country, with the data comprehensive of the insurance and freight (c.i.f.) costs, reported by the importing country, reflects an aggregate measure that refers to all products and depends on the trade composition. Alternatively, other sources[27] have provided data on all those costs related to the movement of goods from the producing factory to the final retail, costs of doing business and logistics services. However, Duval and Utoktham (2011a) underline how these attempts have not been able to capture a detailed measure of trade costs and, although adding them up, they do not provide the real extent of international trade costs.

---

[25] Harmonised standards refer to international standards or to standards related to trading partners (Chen and Novy, 2012).

[26] The c.i.f. acronym stands for Cost, Insurance and Freight and, as in the case of the f.o.b. cost, is mainly intended for maritime shipping. Differently from the f.o.b. cost, the exporter (seller) bears all those costs incurred to get the traded goods to a given point of delivery: insurance, freight charges and all other costs linked to the transit. All other costs to send the goods from the point of delivery to the final destination are paid by the importer (buyer). The difference between f.o.b. and c.i.f. cost lies in the moment/place where the costs and responsibility of the shipment move from the exporter to the importer. With a f.o.b. delivery, costs and responsibility are assumed by the exporter until the origin point (usually a port where goods are moved on board a ship). With a c.i.f., the exporter assumes costs and responsibility until the point of delivery; from the point of delivery to the final destination all charges are borne by the importer. The point of delivery and the final destination could be the same or different.

[27] World Bank Doing Business Report, Logistics Performance Index of the World Bank, Executive Opinion Survey of the World Economic Forum.

The brief discussion since now provided has highlighted more the drawbacks coming from the use of a direct measure of trade costs, rather than the potential advantages it entails. All these limits can be summarised in five main points. First, data availability and measurement problems represent the major problem concerning the use of a direct measure, since just a partial estimate of trade costs is possible. Second, data incompleteness limits the inclusion of some variables and the exclusion of others, creating a potential omitted variables' bias problem.[28] Third, theory does not provide a clear direction for the appropriate trade cost function to be used, generating some arbitrariness among researchers and generating estimated trade cost effects which vary depending on the chosen functional form. Fourth, some trade costs components are time-invariant (like distance, common language, common currency, etc.), and this makes it difficult to take changes in trade costs into account.[29] A further, no less important, limitation connected to the direct measurement lies in the computation issue. The direct method consists in adding up different trade cost sub-components. This way of computing trade costs leads, however, to a typical weakness of economic measurements: the problem of comparing apples and oranges and, more specifically, the problem of adding up apples and oranges. Analogously to GDP, the most comprehensive measure in economics, trade costs involve the collection, the aggregation and the sum of a large amount of data and of vastly different types of elements. These elements are not homogeneous and entail different units of measurement. To find a meaningful measure, the gross domestic product considers the adding up of the market value of single components. For trade costs this is hardly feasible, since it is not possible to express all elements using the same unit of measurement or to convert all of them in monetary terms (i.e. dollars). Expressing trade costs as one-dimensional measure entails a potential bias problem, since trade costs are a composite object with a multiple dimension. Moreover, the possible dependence between some of those variables adds an extra limitation to the computation of trade costs exploiting a direct measurement.

All these drawbacks linked to the direct method have led research to find an alternative way to measure trade costs: to infer them indirectly. Theory has provided two different ways to deduce trade costs. One approach seeks to infer the total barriers to trade from *quantities* (trade flows); the other deduces the extent of trade impediments from *prices*. What is common to both indirect methodologies is that they derive trade costs from observables and provide a comprehensive measure of trade cost not distinguishing among sub-costs

[28] Arvis et al. (2013a).
[29] Chen and Novy (2012).

components. The indirect method is also called '*top-down*' approach, because it infers the level of trade costs from the observed pattern of trade and production. Since the trade costs measure is a function of trade flows and production, the computation requires the use of an economic model to identify the underlying structure of trade costs.

The indirect methodology based on *prices* consists in the idea that arbitrage will eliminate price differences across border.[30] The related literature has been reviewed in the paper of Anderson and van Wincoop (2004). They highlight that inferring trade barriers from final goods prices is more theoretical than practical, since price indices also include non-tradable products and are affected by local taxes, subsidies and nominal rigidities. On that basis, literature has mainly focused on theory and Anderson and van Wincoop observe how two branches can be distinguished. One part of the literature aims to measure trade barriers through the comparison between import or world prices and domestic wholesale prices. The main weakness of this literature is that it measures only a part of the trade barriers (tariff and nontariff barriers, all the costs faced by the importing country), but it doesn't capture all those costs that the exporters bear to carry the good to the foreign market. As observed above, the comparison between prices has been designed to capture nontariff barriers, but, as highlighted by Anderson and van Wincoop, the measurement is not accurate and some weaknesses arise. In addition, incompleteness and imperfection of data limit the comparison of prices. The wholesale price contains some local distribution costs, it usually characterises a domestic substitute of the import good or an index of imported and domestic goods,[31] making them not completely homogenous between countries; moreover, price can be expressed in different currencies, requiring the use of the correct exchange rate. Another strand of the literature aims to measure trade barriers comparing retail prices of goods across countries. The main limit of this literature is represented by the lack of a theory that supports the link between relative prices across countries and trade barriers.

Inferring trade costs from prices neglects the importance of different market structures and the contexts in which exporting firms have market power and are able to set the price according to elasticity of demand (and not only of trade costs). Typically, as pointed out by Anderson and van Wincoop (2004, p. 740) "*the price paid by the final user of a good generally contains four components: (i) the marginal cost of production, (ii) trade costs, (iii) various monopolistic markups over cost in the chain from producer to final user, and (iv) subsidies and taxes*". If the exporter is able to set the import price, then he may keep the revenue for

[30] Novy (2013).
[31] Anderson and van Wincoop (2004).

itself and may affect price differentiation. [32] Price dispersion, as pointed out by Anderson and van Wincoop (2004), does not enable to infer so much about the size of trade barriers, even if trade barriers are the only source of variation and no markups or taxes occur. However, the variability of prices across goods and country pairs allows to obtain a better knowledge about the extent of trade costs.

The indirect methodology based on *quantities* has been widely used, and in literature there are two different ways to infer trade costs from trade flows. One technique, named by Chen and Novy (2011) '*phi-ness measure*', describes a trade costs measure as the ratio of bilateral over domestic trade flows, basically reflecting border-related costs using the gravity model, proposed by Anderson and van Wincoop (2003), as the key framework to compute this measure. In this context, Head and Ries (2001) were the first to use a micro-founded gravity equation. Looking at the U.S.-Canada free trade agreement, they obtain two different measures of trade costs: one measure based on an increasing returns model with home market effects; the second computed using a constant returns model with national product differentiation. Other important contributions are the works of Baldwin et al. (2003) and of Head and Mayer (2004), which compute the phi-ness measure to compare the North-American and European integration. The second technique is the one reported in the pioneering paper of Novy (2013), that represents a milestone in trade costs and international trade literature and that in this thesis is named '*gravity-based measure*'. Novy extends the approach proposed by Head and Ries (2001), deriving from the observed pattern of trade and production a micro-founded, comprehensive and indirect measure of bilateral trade costs using the gravity framework. The novelty of Novy's paper is represented by its theoretical foundation. Novy demonstrates that his trade costs index can be derived from a wide number of typical international trade models: the gravity model by Anderson and van Wincoop (2003), the Ricardian model by Eaton and Kortum (2002), the heterogeneous firms models by Chaney (2008) and Melitz and Ottaviano (2008). As noticed by Arvis et al. (2013a), Novy's approach does not need to assume CES preferences and, since the measure is based on mathematical and theoretical functions, it does not suffer the typical problems of omitted variable or endogeneity bias. The methodology proposed by Novy has been extensively used and the literature provides a wide number of empirical applications.[33] One of the most

[32] Deardorff and Stern (1998).

[33] Jacks et al. (2010), exploiting data on GDP and total trade flows, apply Novy's method to calculate trade costs in the first era of Globalisation for the most important economies; in their contribution of 2008 they compute indirect trade costs according with the gravity-based measurement, to assess whether changes in trade costs can explain trade booms and trade busts from the first wave of Globalisation until today. Arvis et al. (2013a) use the

technical applications of the Novy's methodology is the one proposed by Chen and Novy (2011). Based on the gravity model and on a monopolistic competition framework, the work uses disaggregated trade flows at the industry level and production data categorised by sectors, to compute for 173 manufacturing European industries a measure of trade costs, which accounts for heterogeneity across industries, and for industry-specific substitution elasticities. Differently from the Head and Ries' (2001) method, Chen and Novy (2011) are able to compute and rank trade frictions across countries and industries, since they disentangle the effect of trade barriers and the effect of heterogeneity.

The two ways of measuring trade costs (direct and indirect) are strictly connected to the two ways of modelling trade costs. Some words should be devoted to considering this link. As it will be better clarified in following Paragraph 1.4, differently from the direct approach, the 'top-down' method by Novy (2013) yields an estimate of trade costs that reflects iceberg trade costs. The fact of being based on the gravity model by Anderson and van Wincoop (2003), according to which bilateral trade costs are modelled in ad valorem terms (one plus the tariff equivalent) and increase the price of each unit shipped, makes trade costs à la Novy (2013) an ad valorem equivalent that includes anything that differentiates the price in the exporting country from the price in the importing country.

In conclusion, analogously to what has been argued for the multiplicative-additive issue, both direct and indirect approach are suitable methods to measure trade impediments. The selection of which one to use is strongly related to the investigation that the researcher want to perform. As underlined by Moïsé and Le Bris (2013), both methodologies have strengths and weaknesses. The direct approach has been extensively used in most studies, but it underestimates the complete effect of trade barriers on trade flows, because many of them remain unobservable or data are not available. The indirect approach makes measurements available for many countries and years but is less useful for policymakers, since no distinction within components is possible (making it difficult to target the correct strategies) and since the overestimation of the impact. During these years, literature has gained new contributions, with new insights into the best way to measure trade costs. Anderson and van Wincoop (2004), for the 'bottom-up' method, and Novy (2013), for the 'top-down' method, should be considered as the fathers of the two approaches, but choosing the best one is not possible. On this view, Moïsé and Le Bris (2013) suggest that direct and indirect methodology can be seen as two approaches that complement each other.

---

approach of Novy (2013) to infer estimates of trade costs in agriculture and manufacture sectors in 178 countries for the period 1995-2010. Miroudot et al. (2013) adopt the same approach in services sectors.

1.4 The 'top-down' approach: a gravity-based measure (Novy, 2013)

The 'bottom-up' approach represents the straightforward method to obtain an accurate estimate of trade costs. However, as underlined above, the several drawbacks linked to its use have led research to find an alternative method. Novy in 2013 proposes a gravity-based measure. This innovative methodology is at the core of this thesis: both empirical analyses of Chapter 2 and 3 hang on to the indirect measure of trade costs proposed by Novy (2013). The success of this pioneering approach is highly documented by both theoretical and empirical literature. In this perspective, the current paragraph has been divided into two subparagraphs. Subparagraph 1.4.1 explores the aspects linked to theory: as reminded by the name, the gravity-based measure has a strong theoretical origin. Subparagraph 1.4.2, instead, examines and reviews all those works that put the measure to work.

1.4.1 Theory

The earliest indirect approach to measure trade costs has been developed by Head and Ries (2001), which propose an inverse measure of the 'phi-ness of trade'. To construct this measure, they assume that the intra-national trade is costless, whereas the inter-national trade has symmetric trade costs. Head and Ries (2001) consider two alternative trade models to show how the effect of trade costs changes according to different specialisation and market size features. The first model they consider is the national product differentiation trade model, the Armington[34] (1969) model, which distinguishes industry's goods by nationality, considers a constant elasticity of substitution and includes perfect competition. The second alternative model is the Krugman[35] (1980) model, an increasing returns with imperfect competition (monopolistic competition) model where the variety of products is within firms rather than within countries. The Krugman model is an adaptation of the Dixit-Stiglitz (1977) model with CES (constant elasticity of substitution) preferences. Head and Ries (2001) demonstrate that trade costs reduce the magnitude of the market size effect on output in the increasing returns model, whereas they increase the size of the relationship between the share of output and the share of demand when considering the national product differentiation model. Using data on manufacturing industries in the U.S. and Canada, they find evidence supporting both models, but the reduction in tariff and nontariff barriers is more close to the Armington model.

[34] Armington (1969) suggests that countries produce different goods that are considered by consumers as imperfect substitutes, and they would like to consume at least a part of each country's goods.

[35] Krugman (1980) defines a monopolistic competition model where small firms may costless differentiate their products due their monopoly power.

Starting from the Head and Ries (2001) work, Novy (2013) extends their approach demonstrating that the indirect measure of trade costs originates from different theoretical trade models. In his work he looks at three main models: the gravity model by Anderson and van Wincoop (2003), the Ricardian model by Eaton and Kortum (2002) and the heterogeneous firms models by Chaney (2008) and Melitz and Ottaviano (2008). These models are different in the assumptions and in the explanation of what drives international trade flows, but they all generate gravity equations in general equilibrium. This is important because, since the indirect measure of trade costs is obtained from the gravity equation and since the gravity equation can simply be considered as a consumers' expenditure equation,[36] trade costs can be interpreted as a 'gravity residual' that measures how different existent trade flows are from those predicted by the gravity model in a theoretical frictionless world (Novy, 2013). The indirect approach to compute trade costs shows that the gravity model underestimates trade frictions since it does not consider the domestic trade sector. Trade barriers, in fact, affect not only international exchanges but also internal trade. The insight is that trade flows are affected by a variation in resources between the exporting trade sector and the domestic one; this variation originates, in turn, from a change in trade barriers. Therefore, domestic trade should be included in the gravity equation in order to account for the home bias (Turkson, 2012). Novy (2013), with his measure, attempts to address this need.

The measure proposed by Novy is strictly connected to the so called 'freeness of trade' measure in the New Economic Geography (NEG) field,[37] which quantifies the inverse of trade costs: a high value of the measure means that trade barriers are low; on the other hand, a low value of the freeness of trade corresponds to high trade costs.[38]

The main goal behind the approach of Novy (2013) is to measure the barriers to international trade, solving a gravity equation for the trade cost parameters from different models of trade, in order to obtain a function of observable trade data and a measure of bilateral trade costs.

Novy (2013) proves that the derivation of his gravity-based microfounded trade cost measure can be retrieved from each of the three above-named models. The model of Anderson and van Wincoop (2003) is the first one considered. It is a demand-side multi-country general equilibrium model that describes international trade using a gravity framework and suggests

[36] The gravity equation is basically a consumer's expenditure equation explaining how consumers expend their money across countries, not giving account of the motivation behind their choices (Novy, 2013).

[37] Fujita et al. (1999), Baldwin et al. (2003) and Head and Mayer (2004).

[38] Although Novy does not use any of the traditional models of the NEG literature, he contributes to that literature linking the unobservable multilateral resistance variables to observable data and he demonstrates that trade costs can be derived from some classical (such as the gravity model) and more recent models of international trade.

that, everything else being equal, bigger economies trade more between each other. In the Anderson and van Wincoop's model there is not supply side: production is exogenous and each country is endowed with a single good, which is different from that of the other countries. Individuals diversify their consumption across countries and have identical preferences. As underlined by Novy (2013), the main feature of this model is the introduction of exogenous bilateral iceberg trade costs: trade costs increase the price of each unit of good shipped from country i to country j, making prices different across countries and reducing bilateral trade. In addition, Anderson and van Wincoop use border barrier and geographical distance as a proxy for bilateral trade costs and assume that they are symmetric, implying that the outward and the inward multilateral resistance are the same.[39] The bilateral trade costs measure that Novy obtains from the Anderson and van Wincoop's model is sensitive to the degree of differentiation across products. Differently from Anderson and van Wincoop's model, Eaton and Kortum (2002) propose a Ricardian supply-side general equilibrium model including geography. According to the Eaton and Kortum's (2002) model, the comparative advantage is the driving force behind trade across nations, although geographic barriers represent an important obstacle to integration. In the model, each country can produce all varieties of goods, but just the countries that produce the good at the lowest cost (having an average absolute productivity advantage) ship it to the other countries. Since the comparative advantage drives trade across countries, the different productivity of countries is a fundamental determinant of trade flows. The bilateral trade costs measure derived from the Eaton and Kortum's model depends on the variation of productivity. Both bilateral trade cost measures, derived first from the Anderson and van Wincoop's model and then from the Eaton and Kortum's model, imply that relative trade frictions arise when the heterogeneity (across products or of productivity) is higher, and this is the case when the ratio of domestic over bilateral trade is larger than one. The insight behind the measure is that a higher heterogeneity stimulates trade between countries: if the heterogeneity is high but global trade flows are low, this means that there are trade barriers which impede international trade.

Lastly, Novy (2013) demonstrates how the indirect measure of trade costs can be derived also employing that class of models which focuses on heterogeneous productivity of firms. The model of Chaney (2008) starts from the work of Melitz (2003) and includes heterogeneous

[39] Novy (2013) argues that the assumption of symmetric trade costs is strong and might lead to some pitfalls. First of all, if a country imposes larger tariffs than the country-partner, trade costs are not symmetric anymore. Second, the trade cost function might be misspecified. Third, geographical distance is a weak proxy for trade costs, since it doesn't measure the variation of trade costs over time. In order to solve these drawbacks, Novy (2013) develops an analytical solution for multilateral trade resistance according to which a change in bilateral trade costs has an effect not only on inter-national trade but also on intra-national trade.

firms with a random Pareto productivity distribution and with iceberg trade costs. Chaney proves that the elasticity of substitution affects the extensive margin and the intensive margin of trade, and, differently from Krugman (1980),[40] that the effect of trade barriers on trade flows reduces when the elasticity of substitution is higher. In the Chaney's model, each firm produces a single good and faces bilateral fixed costs to export. The bilateral measure of trade costs Novy obtains is sensitive to the entrance and the exit of the firms into and from foreign markets. The model of Meltiz and Ottaviano (2008) is a monopolistic competitive model with firm heterogeneity in terms of productivity differences. Market size and trade affect competition in a market, leading to the selection of heterogeneous exporting firms in that market. They outline a model where heterogeneous firms face variable costs of exporting. The bilateral trade cost measure Novy derives from this model, as in the measure derived from Chaney's model, depends on the entrance and the exit of the firms into and from foreign markets, although in this case fixed costs are not included, since firms face only variable costs of exporting.

To better appreciate the insights of the gravity-based measure proposed by Novy (2013), it is not possible to disregard the mathematical derivation of the trade cost equation from the gravity model by Anderson and van Wincoop (2003) that Novy provides in his contribution. Here a very detailed derivation is presented in order to allow a full understanding of the index.[41]

Anderson and van Wincoop (2003) define the final equation representing the gravity model as follows:

$$x_{ij} = \frac{y_i y_j}{y^W}\left(\frac{t_{ij}}{\Pi_i P_j}\right)^{1-\sigma} \qquad (1)$$

Where:

- x_{ij} represents the nominal value of exports from i to j (j's payments to i)
- y_i represents the nominal income of country i
- y_j represents the nominal income of country j
- y^W represents the world income
- t_{ij} represents the bilateral trade barriers between i and j

[40] Krugman (1980), looking at identical firms, explains that the impact produced by trade barriers on international exchanges is higher when the elasticity of substitution between goods increases.

[41] Arvis et al. (2013a), Duval and Utoktham (2011b), Turkson (2012) and Gaurav and Mathur (2016) provide in their works a similar derivation.

- Π_i represents the (outward) multilateral resistance of country i (exporter)
- P_j represents the (inward) multilateral resistance of country j (importer)
- σ represents the elasticity of substitution between all goods

According to Anderson and van Wincoop (2003), bilateral trade between country i and country j depends on the product of their incomes relative to the world income and on the bilateral trade barriers relative to the product of the two multilateral resistance terms. They assume that bilateral trade costs are symmetric (i.e., $t_{ij} = t_{ji}$). This generates equal outward and inward multilateral resistance terms (i.e. $\Pi_i = P_j$). Novy highlights how this assumption has some weaknesses, and, in order to overcome these drawbacks, he derives an analytical solution for multilateral resistance variables, making use of the insight that a change in bilateral trade barriers affects not only inter-national trade but also intra-national trade. The solution for multilateral resistance variables is useful to solve the model for bilateral trade costs. Using country i as reference country,[42] the gravity equation can be rewritten in terms of country i's intra-national trade:

$$x_{ii} = \frac{y_i y_i}{y^W} \left(\frac{t_{ii}}{\Pi_i P_i}\right)^{1-\sigma} \quad (2)$$

t_{ii} becomes country i's intra-national trade costs
Solving for outward and inward multilateral resistance terms:

$$x_{ii} = \frac{y_i y_i}{y^W} \left(\frac{\Pi_i P_i}{t_{ii}}\right)^{\sigma-1}$$

$$(\Pi_i P_i)^{\sigma-1} = \frac{x_{ii} y^W}{y_i y_i} t_{ii}^{\sigma-1}$$

$$\Pi_i P_i = \left(\frac{x_{ii} y^W}{y_i y_i}\right)^{\frac{1}{\sigma-1}} t_{ii} \quad (3)$$

[42] The same applies for country j.

Equation (3) implies that for given domestic trade costs t_{ii} it is possible to calculate the change in multilateral resistance over time, as it does not depend on time-invariant trade cost proxies such as distance. In fact, "*if two countries i and j, have the same domestic trade costs $t_{ii} = t_{jj}$ and are of the same size $y_i = y_j$ but country i is a more closed economy, that is, $x_{ii} > x_{jj}$, [...] the multilateral resistance is higher for country i ($\Pi_i P_i > \Pi_j P_j$)*" (Novy, 2013, p. 105).

The equation (1) contains the product of outward multilateral resistance of one country and inward multilateral resistance of another country, $\Pi_i P_j$. The equation (2) represents the solution for $\Pi_i P_i$. In order to have a bidirectional gravity equation that contains both countries' outward and inward multilateral resistance variables, equation (1) needs to be multiplied by the gravity equation for trade flows in the opposite direction (x_{ji}):

$$x_{ji} = \frac{y_j y_i}{y^W} \left(\frac{t_{ji}}{\Pi_j P_i} \right)^{1-\sigma} \quad (4)$$

Equation (4) represents the exports from j to i (x_{ji}).

Multiplying (1) by (4) yields:

$$x_{ij} x_{ji} = \left(\frac{y_i y_j}{y^W} \right)^2 \left(\frac{t_{ij} t_{ji}}{\Pi_i P_i \Pi_j P_j} \right)^{1-\sigma} \quad (5)$$

Rearranging, equation (5) can be rewritten as:

$$x_{ij} x_{ji} = \left(\frac{y_i y_j}{y^W} \right)^2 \left(\frac{\Pi_i P_i \Pi_j P_j}{t_{ij} t_{ji}} \right)^{\sigma-1}$$

Substituting the solution from equation (3) and rearranging yields:

$$x_{ij} x_{ji} = \left(\frac{y_i y_j}{y^W} \right)^2 \left(\frac{1}{t_{ij} t_{ji}} \right)^{\sigma-1} \left(\frac{x_{ii} y^W}{y_i y_i} \right)^{\frac{\sigma-1}{\sigma-1}} t_{ii}{}^{\sigma-1} \left(\frac{x_{jj} y^W}{y_j y_j} \right)^{\frac{\sigma-1}{\sigma-1}} t_{jj}{}^{\sigma-1}$$

$$x_{ij} x_{ji} = \left(\frac{y_i y_j}{y^W} \right)^2 \left(\frac{1}{t_{ij} t_{ji}} \right)^{\sigma-1} \left(\frac{x_{ii} y^W}{y_i y_i} \right) t_{ii}^{\sigma-1} \left(\frac{x_{jj} y^W}{y_j y_j} \right) t_{jj}^{\sigma-1}$$

$$x_{ij} x_{ji} = \left(\frac{t_{ii} t_{jj}}{t_{ij} t_{ji}} \right)^{\sigma-1} x_{ii} x_{jj}$$

$$\frac{x_{ij} x_{ji}}{x_{ii} x_{jj}} = \left(\frac{t_{ii} t_{jj}}{t_{ij} t_{ji}} \right)^{\sigma-1}$$

$$\left(\frac{t_{ij} t_{ji}}{t_{ii} t_{jj}} \right)^{\sigma-1} = \frac{x_{ii} x_{jj}}{x_{ij} x_{ji}}$$

$$\frac{t_{ij}t_{ji}}{t_{ii}t_{jj}} = \left(\frac{x_{ii}x_{jj}}{x_{ij}x_{ji}}\right)^{\frac{1}{\sigma-1}} \quad (6)$$

Equation (6) says that the product of bilateral trade costs (t_{ij} * t_{ji}) relative to the product of intra-national trade costs (t_{ii} * t_{jj}) corresponds to the product of internal trade (x_{ii} * x_{jj}) relative to the product of bilateral trade (x_{ij} * x_{ji}) to the power of 1/(σ-1). Since both bilateral and domestic trade costs can be asymmetric (i.e. $t_{ij} \neq t_{ji}$ and $t_{ii} \neq t_{jj}$), it is helpful to take the geometric mean of the bilateral trade costs:

$$T_{ij} = \left(\frac{t_{ij}t_{ji}}{t_{ii}t_{jj}}\right)^{\frac{1}{2}} = \left(\frac{x_{ii}x_{jj}}{x_{ij}x_{ji}}\right)^{\frac{1}{2(\sigma-1)}} \quad (7)$$

The tariff equivalent total trade costs (τ_{ij}) could be obtained by subtracting one in both directions:

$$\tau_{ij} = \left(\frac{t_{ij}t_{ji}}{t_{ii}t_{jj}}\right)^{\frac{1}{2}} - 1 = \left(\frac{x_{ii}x_{jj}}{x_{ij}x_{ji}}\right)^{\frac{1}{2(\sigma-1)}} - 1 \quad (8)$$

where:[43]

- τ_{ij} represents geometric average trade costs between country i and country j
- t_{ij} represents international trade costs from country i to country j
- t_{ji} represents international trade costs from country j to country i
- t_{ii} represents intra-national trade costs of country i
- t_{jj} represents intra-national trade costs of country j
- x_{ij} represents exports from country i to country j
- x_{ji} represents exports from country j to country i
- x_{ii} represents intra-national trade of country i calculated as GDP minus total exports[44]
- x_{jj} represents intra-national trade of country j calculated as GDP minus total exports[45]
- σ represents the elasticity of substitution across goods, where $\sigma>1$.

Equation (8) is the final bilateral comprehensive trade cost measure defined by Novy (2013) and states that the total trade cost (τ_{ij}) corresponds to the product of bilateral trade costs (t_{ij} *

[43] Duval and Utoktham (2011a, 2011b) provide a similar description.
[44] x_{ii} may also be named 'internal trade' of country i.
[45] x_{jj} may also be named 'internal trade' of country j.

t_{ji}) relative to domestic trade costs (t_{ii} * t_{jj}). In other terms, it represents the geometric average of international trade costs between country i and country j relative to domestic trade costs within each trade partner. According to the measure, when the ratio rises,[46] trade costs are higher and countries are more likely to trade domestically than internationally. On the contrary, as the ratio falls,[47] trade costs are lower and countries tend to trade more with their trading partners than they do with themselves. In other words, data on the relative openness of a country can be interpreted as the extent of its trade costs: if the country ships abroad the part of its production that was previously consumed domestically, it means that the country is more open and its trade costs are lower. Since trade costs originate from a ratio with trade flows in the denominator, country pairs that do not trade at all record infinite trade costs.[48] Moreover, the measure allows for asymmetric bilateral trade costs ($t_{ij} \neq t_{ji}$) and for unbalanced trade flows ($x_{ij} \neq x_{ji}$) between the pair. "*It includes all factors that contribute to the standard definition of iceberg trade costs in trade models, namely anything that drives a wedge between the producer price in the exporting country and the consumer price in the importing country. Trade costs [...] therefore include both observable and unobservable factors*" (Arvis et al., 2013b, p. 11).

The value of the elasticity of substitution affects the indirect measure of trade costs. However, Arvis et al. (2013a) suggest that results are not sensitive to the used parameter. Novy (2013) assumes that the elasticity is constant across sectors, countries, and years and sets an elasticity of substitution equal to 8. In literature various approaches and different estimates have been proposed; Paragraph 1.6 gives a brief review of the three main methods to calculate σ and the main values suggested.

The way according to which the indirect measure of trade costs is constructed involves both advantages and drawbacks and needs to be taken into account to optimise its empirical application. Both potentials and pitfalls of the indirect approach could be viewed, respectively, as disadvantages and advantages of the direct method. The first strength of the 'top-down' method lies in the ease of the empirical implementation with country level and long series available data, making it extensively useful in studies with numerous countries[49]

[46] International trade costs rise relative to domestic trade costs.
[47] International trade costs fall relative to domestic trade costs.
[48] Arvis et al. (2013a).
[49] Arvis et al. (2013a) compute trade costs according with the indirect method proposed by Novy (2013) for a wide number (178) of developing and developed countries.

and long term period.[50] One of the main problems in implementing the 'bottom-up' approach is, instead, the poor coverage of data in terms of countries and years: as underlined by Anderson and van Wincoop (2003), data on tariff and nontariff measures are incomplete and just major economies (mostly OECD countries) or recent years include an exhaustive reporting. The second advantage of Novy's approach is the 'all inclusive' character of the measure, comprehensive of all those components of trade costs that are difficult to observe or to measure using a direct method (like language barriers, informational costs, exchange rate risk, nontariff barriers etc.). Third, differently from every gravity approach, where a function is needed to rely on some geographical explanatory variables (such as distance), the indirect measure infers trade costs from observable data, not requiring the use of a trade cost function. The existence of some time-invariant elements (like distance or language) would lead to a static trade costs function, which imperfectly captures the variation of trade costs over time, but the Novy's measure is a function of time-varying bilateral trade data generating a time-varying bilateral trade cost measure and producing a further advantage for researchers, since it enables to assess changes in trade costs over time. The fifth strength of this indirect approach is the absence of a possible problem of omitted variables or endogeneity bias: since it is a theory-based measure and not an econometric estimation, there is no chance to omit important variables from the measurement or to run into simultaneity. On the contrary, the direct approach, due the mathematical addition nature, is seriously affected by the bias of omitted covariates.

This measure is also plagued by some drawback, hence it needs to be interpreted cautiously. First, since the measure "*is the geometric average of trade costs in both directions, [...] from a policy perspective, it is therefore impossible to say without further analysis whether a change in trade costs between two countries is due to actions taken by one government or the other, or both together*". For this reason, "*further analysis is required [...] to identify the sources of trade costs and their relative contributions to the overall number*" (Arvis et al, 2013a, p. 11). Single sub-components (like freight rates) represent just a fraction of overall trade costs. In these terms, the comprehensive nature of the indirect measure of trade costs emerges since only a non-identifiable part of the total estimate is due to direct policy interventions by governments.

[50] Jacks et al. (2008, 2010 and 2011) apply the top-down approach computing trade costs for the first era of Globalisation (43 years, from 1870 to 1913) and for the entire period, between 1870 and 2000 (130 years), including the first and the second (1950-2000) wave of globalisation period and the interwar period.

Second, the interpretation of τ_{ij} depends on the theoretical model from which it is derived: in the gravity model by Anderson and van Wincoop (2003) trade costs are variable and can be interpreted as having an ad valorem structure; in other models with fixed costs, such as in Chaney (2008), trade costs are a combination of both fixed and variable elements (Arvis et. al, 2013a). A third pitfall lies in the precision of the indirect approach: due the construction using not true components of trade costs, the accuracy of the measure may be debated;[51] on the other hand, the direct measure has the advantage of being more accurate, since it includes 'real' trade costs sub-components, and, therefore, allows to disentangle what is the contribution that each trade cost element adds to the total extent of trade costs. The indirect approach does not allow a similar inference, and just an aggregated and general understanding of trade costs is possible, without a detailed focus on the single factors. The fourth limitation has to do, instead, with price effects. Arvis et al. (2013a) observe that, since Novy nets multilateral resistance out from the measure, then price changes are also eliminated from the approach. In fact, the outward and inward multilateral resistance variables Π_i and P_j are respectively country i's and country j's price indices, that represent deflators for GDP and trade values, and bilateral trade costs t_{ij} are a measure against these price indices. By finding an explicit solution for the multilateral resistance terms, Novy is able to define trade costs τ_{ij} without those variables. In light of the fact that trade values may change at a different rate from total production values, and if only goods having very high value to weight ratios are traded, then τ_{ij} should be interpreted carefully because of the potential to join unit price and volume effects.

1.4.2 Empirics

When taking the trade cost measure to data, it is possible to compare what different studies find using the same approach. Putting the measure to work is fundamental to understand whether and how much the measure is reliable. Duval and Utoktham (2011a) underline that the comparison between estimates computed using the same methodology is appropriate. On the other hand, evaluating approximations obtained from different approaches is unsuitable and wrong. Nevertheless, it allows to assess whether different methodologies lead to similar results. Anderson and van Wincoop (2004) suggest that an approximation of overall ad valorem trade costs for developed countries is 170 percent, divided into 55 percent local distribution costs and 74 percent international trade costs. This is the most influential

---

[51] See Paragraph 1.5 for more information.

estimation in the literature and, although it comes from a 'bottom-up' approach, it can be used as a benchmark. Using the 'top-down' approach, Novy (2013), Shepherd (2010), Duval and Utoktham (2011a), Arvis et al. (2013a), and Turkson (2012) provide the most significant contributions in estimating current trade costs for a different set of countries. All these works assume an elasticity of substitution constant over the time, and set it equal to 8. Table 1.1 summarises all these findings.

Table 1.1 Main contributions measuring trade costs according to the indirect approach by Novy (2013)

Author(s)	Aim of the work	Observed countries	Observed time span	σ	Main evidence on trade costs
Novy (2013)	To measure barriers of international trade by developing a micro-founded measure of aggregate bilateral trade costs from the gravity equation	The U.S. (with its main six trading partners: Canada, Germany, Japan, Korea, Mexico, the U.K.)	1970-2000	8	1) Decrease of trade costs of 40% over the period 1970-2000; 1.9% per year 2) With Mexico and Canada (biggest trading partners) decrease reaches respectively 66% and 50%
Shepherd (2010)	To understand trade facilitation in terms of trade costs reduction	APEC countries ASEAN countries	1995-2008 2001-2007	8	1) APEC countries: decrease of trade costs of 5% over the period 2001-2006; for biggest APEC countries the fall reaches 15% 2) ASEAN countries: decrease of trade costs is lower than for APEC countries and it is, in GDP-weighted terms, of 2%
Duval and Utoktham (2011a, 2011b)	To evaluate intraregional trade costs within and among Asian sub-regions, and extraregional trade costs between Asia and foreign regions	ASEAN East and North-East Asia North and Central Asia SAARC EU5 NAFTA Australia-New Zealand MERCOSUR	2003-2007	8	1) Intraregional trade costs lower than interregional ones 2) Extraregional trade costs lower that interregional ones (with some exceptions 3) Reductions of trade costs are more consistent in trade outside Asia
Arvis et al. (2013a)	To infer trade costs for a wide group of countries distinguishing by WB income category and by agriculture and manufacturing sector	178 developed and developing countries	1995-2010	8	1) Trade costs for rich countries lower than for poor countries. 2) Trade costs for agriculture sector higher than for manufacturing sector 3) Reduction of trade costs for the manufacturing sector over the period 1996-2009: 15% for high income group, roughly 5% for low income category
Turkson (2012)	To estimate trade costs for the sub-Saharan Africa	34 sub-Saharan African countries	1980-2003	8	1) Sub-Saharan Africa faces the highest trade costs in the world 2) Increase of trade costs over the period 1980-2003 only for sub-Saharan countries

Source: Author's elaboration

The pioneering work of Novy (2013), besides the novelty of a micro-founded measure of aggregate bilateral trade costs, also includes an empirical part where the author provides a measure for a number of leading countries in international trade. He computes bilateral trade costs for the U.S. with its main six trading partners (Canada, Germany, Japan, Korea, Mexico and the U.K.) over a thirty years period (1970-2000). Novy finds that, in 1970, the measure of bilateral trade costs relative to domestic ones ranged from 50 percent (between the U.S. and Canada) to 107 percent (when looking at the U.S.-Korea trade). In 2000, the level of trade costs ranks from 25 percent (between the U.S. and Canada) to 70 percent (when considering

Germany or Korea).[52] In thirty years trade costs experienced a decrease of 40 percent.[53] Novy's results are consistent with those found by Anderson and van Wincoop (2004) and with what the literature of trade costs asserts about the positive effect of improvements in transportation and in technology in lowering trade costs, and confirm that, being in a common trade union (such as the NAFTA[54]), reduces the barriers between countries.

Shepherd (2010) tries to assess the extent of trade costs for the APEC[55] members between 1995 and 2008 and for the ASEAN[56] countries between 2001 and 2007. He obtains that for APEC states the amount of trade costs in 2006 ranges from a minimum of 35 percent (between China and the rest of the world) to a maximum of 86 percent (between Brunei Darussalam and the rest of the world); for some APEC countries (such as China, Peru and Viet Nam) trade costs fall of 15 percent from 2001. When considering all the APEC countries, the reduction of trade costs is, in simple average terms, of 5 percent. For the ASEAN group in 2003 (year with the biggest number of observations) trade costs vary from 22 percent (between Malaysia and rest of the world) to 92 percent (between Brunei Darussalam and the rest of the world). When considering trade between the World and all the APEC or between the World and all the ASEAN members, trade costs in 2006 are respectively 56 percent and 53 percent. Reductions of trade costs for ASEAN countries are lower compared to those of APEC countries: in GDP-weighted terms, trade costs fell from 57 percent to 55 percent in the period 2001-2007. These values are hard to compare with those of Novy or with the evidence coming from the literature, since in Shepherd (2010) the bilateral trade costs refer to a country i (one member of the APEC or ASEAN organisation) and a country j that is not namely a country, but the rest of the World.

Duval and Utoktham (2011a) compute bilateral trade costs from 2003 to 2007 for eight different regions[57] of countries, including overall 31 different states. A first result refers the comparison between intraregional and interregional trade costs: Duval and Utoktham find that the first ones are higher than the second ones, reasonably due to distance not well addressed

---

[52] Bilateral trade costs between the U.S. and Japan are 65 percent, 60 percent, when considering the U.K, and 33 percent with Mexico (Novy, 2013).

[53] 38 percent if computing a simple average, 44 percent in terms of a trade-weighted average (Novy, 2013).

[54] NAFTA is the acronym of North American Free Trade Agreement, that establishes a trade bloc between the U.S., Canada and Mexico.

[55] APEC stands for Asia-Pacific Economic Cooperation, a forum of 21countries (many of them with a coastline in the Pacific Ocean) that promotes free trade in the Asia-Pacific region.

[56] ASEAN stands for Association of South-East Asian Nations, a regional organisation composed by the 10 Southeast Asian countries that promotes intergovernmental cooperation and economic integration amongst its members.

[57] The eight groups are ASEAN, East and North-East Asia, North and Central Asia, SAARC, EU5, NAFTA, Australia-New Zealand, MERCOSUR.

by policy measures. A second evidence comes from the comparison of interregional trade costs between the four Asian sub-regions with the extraregional trade costs between these four Asian groups and four foreign regions (EU5, NAFTA, Australia-New Zealand, MERCOSUR). It emerges that interregional trade costs are on average higher, although some exceptions for the East and North-East Asia and the ASEAN groups occur. Indeed, when looking at the Asian sub-regional trade costs, values range from 61 percent (between ASEAN countries) to 359 percent (when considering trade between Australia-New Zealand and the North and Central Asia). These high values are quite reasonable, since the majority of countries included in the analysis are developing countries. Nevertheless, the findings are in line with those coming from prior ad hoc studies on Asian countries. Moreover, the reduction of trade costs has been achieved mainly in trade with foreign developed countries than within Asia. A further result of Duval and Utoktham's contribution refers to the trade costs between the U.S. and its two main trading partners. The estimate, although unrelated to the intentions of the work, is intended to be compared to estimates coming from similar studies. Duval and Utoktham find that trade costs in 2007 between the U.S. and Canada are 41 percent, when considering the U.S. and Mexico 47 percent. These calculations, as highlighted by the authors, are consistent with those computed by Novy using GDP data.

Arvis et al. (2013a) look at 178 developed and developing countries in the period 1995-2010, classifying states into World Bank income groups and regions and distinguishing between trade in manufactured products and trade in agricultural goods. As extensively highlighted by the literature, trade costs for rich countries are lower than those for poor countries: Arvis and co-authors estimate that in 2009 trade costs for high income and low income countries are an average ad valorem equivalent of 115 percent and 275 percent, respectively, in the manufacturing sector. Reductions in trade costs occur for all groups of countries between 1996 and 2009; however, in the high income group the fall is higher (15 percent) compared to the one in the low income category (5 percent). A second result attains the difference between manufacturing and agriculture sectors. Trade costs for the agricultural goods are higher. Within the group of developing countries they are twice (246 percent) the trade costs for manufactured products. For developed countries this difference is not so high: trade costs for the agriculture sector are 20 percent higher than those for the manufacturing sector. When looking at trade between WB income categories, Arvis and co-authors find that trade costs between a pair of both high income countries are 130 percent. This estimate rises when considering dissimilar countries: between a high income country and a low income country the amount of trade costs is 289 percent. But when considering two low income countries, the estimate decreases to 217 percent. These results suggest that trade is less costly in rich countries and among similar-income countries. When considering interregional trade, trade costs among the East Asia and Pacific region are quite in line with the Duval and Utoktham's (2011a) and Shepherd's (2010) findings.

The last work considered is the contribution of Turkson (2012), where the trade between sub-Saharan African states and foreign countries is analysed in order to estimate the bilateral measure of trade costs. Turkson groups all 155 countries in the sample in eight regional-economic categories (sub-Saharan Africa, EU, North America, East Asia and Pacific, rest of Europe and Central Asia, Latin America and the Caribbean, Middle East and North Africa, South Asia) to facilitate comparisons. Trade costs for the sub-Saharan Africa group are the highest compared to the other groups; they reach an average tariff equivalent of 272 percent. Moreover, whereas the other countries experience a reduction over the period 1980-2003, this is not the case for sub-Saharan African countries, for which trade costs experience an increase. The work of Turkson, although focused on the sub-Saharan countries, allows to say something more about the other countries. During the five years between 1980 and 1984 the EU and the North America countries experience the lowest trade costs, with an average ad valorem of 197 percent. The Latin America and the Caribbean, instead, report the highest: 270 percent. In roughly twenty years the picture has not changed. North America has the lowest trade costs (175 percent) in the period 2000-2003, followed by the EU (187 percent) and South Asia (208 percent).

1.5 Indirect trade cost measure: (new) thoughts

The present paragraph aims to provide more insights on the indirect measure of trade costs. The way according to which the measure is constructed may give rise to some doubts on what the measure actually captures, and leads to some thoughts on its use at a more disaggregated level than at the country one.

The main critique on the indirect measure of trade costs lies in its capability to capture the true frictions of trade. The elements composing the measure are essentially three: GDP, total exports and bilateral exports. None of these components has directly to do with trade barriers or with costs, and the doubt is whether the measure is really assessing the extent of barriers rather than being simply related with GDP or inversely correlated with exports. The question is basically whether changes in GDP or variations in exports are reflecting changes in trade costs or are simply the consequence of other causes completely unrelated with them. GDP is a proxy for the total value of production. Differently, trade is not based on value added, but is expressed as gross exports.

Some extra words about the use of proxies in the computation of trade costs, according to the indirect approach, are needed. Novy (2013, p. 107) highlights that "*[...] gross domestic product (GDP) data are not suitable as income y_i because they are based on value added, whereas the trade data are reported as gross shipments*". Since trade is expressed as gross exports, when computing domestic trade, the ideal measurement requires to subtract exports from gross output. Unfortunately, gross output data is not available for most of the countries and GDP is used as a substitute. Since the unavailability of the gross output information, part of the literature, like Duval and Utoktham (2011a, 2011b), Shepherd (2010), Turkson (2012) exploit GDP as income measure. Arvis et al. (2013a), instead, describe the big work they do to obtain, on the one hand, gross domestic shipments data and, on the other hand, gross output data. As highlighted by Arvis and co-authors, value added data understate the value of production, since they exclude intermediate goods. When considering internal trade, value added has an opposite effect. As pointed out by Novy (2013, p. 109), "*GDP data tend to overstate the extent of intranational trade and thus the level of trade costs because they include services*". In fact, when comparing bilateral trade cost measure between the U.S. and Canada in 1993, the difference between Anderson and van Wincoop's (2004) estimate and the Novy's one is ascribed to the different income measure they use: Anderson and van Wincoop (2004) exploit GDP, Novy adopts merchandise production, following Wei (1996). However, Novy (2013), Jacks et al. (2011), Duval and Utoktham (2011a, 2011b) and Shepherd (2010) underline that there are no significant differences in percentage changes of trade costs between GDP and gross output over time. GDP and gross output are highly correlated, making GDP a reliable proxy for the total value of production.

A second critique lies in the productivity of firms. A high value of the indirect trade cost measure is interpreted, according to Novy (2013), as intra-national trade; this is less costly than inter-national trade and, therefore, countries prefer to trade more domestically than with foreign partners. However, a high value of trade costs may hide a problem of firms' productivity: countries don't export to foreign countries not because trade costs are high, but because their firms have little to export, are not so productive or cannot afford the fixed cost to enter into exporting. This critique is grounded in the work of Melitz (2003). Melitz demonstrates that, given the assumption of fixed costs to enter into a new foreign market and given heterogeneous productivity among firms, there exists a productivity threshold that firms must exceed in order to export, which results in the extensive and intensive margin to trade.

This deviates from what a part of the gravity-type models says.[58] Anderson and van Wincoop (2003) argue that, given products heterogeneity between countries, consumers' love-for-variety and homogeneous firms with identical productivities, there is no correction in the extensive margin, but just in the intensive one. The indirect measure of trade costs by Novy (2013) might capture exactly this aspect: countries don't trade not because the cost of trade is high, but just because the firms are not able to export. On the other hand, Chaney (2008) extends Melitz's model and proves that the elasticity of substitution between product varieties affects barriers to trade, which in turn influence both intensive and extensive margins. The extensive margin is clearly more sensitive to trade barriers than the intensive one. When the elasticity of substitution is low, products are more heterogeneous and less substitutable. In this case, trade barriers are high and firms don't need to compete. Differently, when the elasticity of substitution is high, products are more comparable and more substitutable, trade barriers are low and less productive firms can enter the market. From this perspective, the work of Chaney allows to reconsider the doubt in light of the fact that a higher indirect measure of trade costs implies higher trade barriers and less competition among firms which are not tempted to enter in new foreign markets.[59] Moreover, it can be argued that, if firms don't export, it is because they have in general high trade costs, disregarding where the high trade costs come from: if they are due to the high transport costs, for instance, or because the firms are less productive.

The third aspect that might be criticised concerns the application of the measure at a disaggregated level (NUTS2 or NUTS3[60]) rather than at the country one. Capturing trade costs at the regional or provincial level may be complex using an indirect measure in the way constructed by Novy (2013), since some limits and weaknesses may arise. When one of the trading partners is represented by a region or a province,[61] the computation of trade costs regarding the bilateral region-country or province-country trade might be imprecise, due the inaccuracy in the calculation of the regional or provincial internal trade. In Paragraph 1.4, internal trade has been defined as the total production of the trading partner minus its total

---

[58] This is the literature without heterogeneous firm framework, that precedes the contributions of Chaney (2008), Melitz and Ottaviano (2008), and Helpman, Melitz and Rubinstein (2008).

[59] In Melitz and Ottaviano (2008) heterogeneous firms only face variable costs; fixed costs to enter in a foreign market can be interpreted as "*product development and production start-up costs*" and are incurred before the entry in a new exporting market.

[60] The NUTS classification (Nomenclature of territorial units for statistics) is a hierarchical system for dividing up the economic territory of the European Union. NUTS1 corresponds to the major socio-economic regions; NUTS2 corresponds to basic regions for the application of regional policies; NUTS3 corresponds to small regions for specific diagnoses (Source: http://ec.europa.eu/eurostat/web/nuts/overview).

[61] This is the case when i=region and j=country or i=province and j=country.

exports. According to this definition, to calculate the internal trade of a region or a province, the researcher should subtract from the regional or provincial GDP the total regional or provincial exports. However, this is not enough, because the researcher should also consider the regional or provincial exports that the region or the province carries out outside the regional or provincial border but inside the home country, i.e. the exports to other regions or other provinces of its home country. Unfortunately, ordinary data on total regional or provincial exports of national statistics offices includes only the total exports that the region or the province performs outside the country in which it is included. A precise internal trade at sub-country level should be, instead, computed as follows:

$$Internal\ Trade_i = GDP_i - Total\ Exports_{i \rightarrow countries} - Total\ Exports_{i \rightarrow home\ country}$$

This is less the case when considering interregional or interprovincial trade, i.e. bilateral trade between trading partners that are both regions or both provinces. In this case, total exports should be computed including all exports to the other regions or provinces. For this kind of analysis the main difficulty is represented by the data availability. There are at the moment no data on interprovincial trade or on world interregional (NUTS2 level) trade.[62]

Another imprecision may lie in the regional or provincial GDP, due some ‘external factors’ that may contaminate the exact calculation. As well known, the GDP is simply the adding up of consumption (C), investments (I), public expenditure (G), exports (X) and imports (M) and can be expressed by the following equation:

$$GDP = C + I + G + X - M$$

Considering consumption, families don't consume only goods produced inside the region or the province but also goods produced outside the regional or provincial boundary. The same can be argued for the investments. From this perspective, when computing regional or provincial GDP, to obtain an exact approximation, the researcher should correct the measure subtracting all those ‘influences’ coming from outside the regional or provincial border. This kind of correction is not feasible, due the lack of similar information.

[62] However, for the 276 EU regions, the Joint Research Centre of the European Commission provides for three selected years (2000, 2005, 2010) the data about the trade between the European regions. Future research will exploit these data to investigate interregional trade costs.

The brief discussion here provided has highlighted some weaknesses linked to the indirect way to compute trade costs. Similarly to other measures in economics, the ability to fully and precisely capture an economic phenomenon is hard, since, beyond the advantages connected to the use of a specific method, limitations arise as well. This thesis and, more generally, current economic research, due missing data, cannot mitigate the highlighted drawbacks. However, the thoughts proposed in this paragraph should be interpreted as a preliminary step and further discussion is needed in order to provide a more detailed understanding. Moreover, the majority of constraints linked to the use of an indirect approach depends on data availability. Therefore, the effort devoted to solving or mitigating the weaknesses of the measure should take into account this big burden. Although all the critiques that can arise and although the current incapability to overcome all limitedness of the measure, it should be underlined that the potentials of the indirect approach are large and allow to perform consistent empirical analyses, relying on a strong method that originates from the theory. The strength of the measure lies exactly in the way in which it is constructed, that leaves space to adaptations and to tailored revisions according to the needs of the researcher.

1.6 The elasticity of substitution in the indirect trade costs measure

According to Anderson and van Wincoop (2004), the elasticity of substitution used affects the value of trade costs. If considering more aggregate data or adopting some more industry-specific focus, the choice of the correct parameter might have a different relevance. The literature proposes a range of values among which the indirect measure of trade costs seems to be not sensitive to the used parameter. Novy (2013) uses a constant elasticity of substitution across sectors, countries and years. On the other hand, Chen and Novy (2011) demonstrate that the elasticity of substitution varies across industries and depends on technology parameters.

Anderson and van Wincoop (2004) describe the three different ways used in literature to obtain an estimate of σ: one method comes from the observation of trade barriers; a second approach refers to demand equations; a third method is the approach proposed by Eaton and Kortum (2002).

The first procedure to obtain σ refers to directly observed trade barriers, like tariffs and transport costs, and to estimating gravity equations using sectoral data for two or more countries. As suggested by Anderson and van Wincoop, the most relevant contributions to this approach are those of Hummels (2001), Head and Ries (2001), Baier and Bergstrand

(2001) and Harrigan (1993). Hummels (2001), considering a tariff rate and freight factor and using data on sectoral imports of six countries from a large number of other nations, finds that the elasticity increases from 4.8 for one-digit SITC data to 8.3 for four-digit SITC data. Head and Ries (2001) consider only two countries, the U.S. and Canada, three-digit industry data, and tariff and nontariff barriers. They obtain an elasticity of substitution of 11.4 when nontariff components don't vary across industries, and of 7.9 when changes are allowed. Baier and Bergstrand (2001), referring to OECD countries and using only tariffs and transport costs as trade barriers, obtain 6.4 as an estimate of the elasticity of substitution. Harrigan (1993) estimates the effects of nontariff barriers on the bilateral imports of ten large OECD countries from 13 trading partners in 1983 for 28 sectors, obtaining different estimates of σ, that ranges from 5 to 10.

The second method obtains estimates for σ from demand equations, using data on price. Anderson and van Wincoop indicate the contribution of Feenstra (1994) to this approach. In its work Feenstra uses the fact that variance and covariance of demand and supply changes have a linear relationship depending on demand and supply elasticities. He computes disaggregated elasticities of substitution using data on imported quantities and prices. Using data on U.S. imports for six manufacturing sectors with more than eight-digit SITC, the estimates for the elasticity of substitution range from 3 to 8.4. The method proposed by Feenstra (1994) has been recently adapted by Broda and Weinstein (2006), Imbs and Méjean (2009) and Chend and Novy (2011).

Eaton and Kortum (2002), employing a specific equation to calculate σ, use data on retail price levels for fifty manufactured products in nineteen countries, obtaining an estimate of 9.28.

From these three different ways to compute the elasticity of substitution of trade costs, Anderson and van Wincoop conclude that σ is likely to range from 5 to 10 and they suggest that estimates for goods that are more differentiated and therefore less substitutable, are around 7 or 8. Arvis et al. (2013a) highlight how the choice of larger values is mainly an issue of assumption rather than measurement. As pointed out in Subparagraph 1.4.2, Novy (2013), in the empirical part of his paper, uses $\sigma=8$, that is a suitable value for aggregate trade flows. Accordingly, Shepherd (2010), Duval and Utoktham (2011b), Arvis et al (2013a), Turkson (2012) and Miroudot et al. (2013) follow Anderson and van Wincoop (2004) and Novy (2013) and set $\sigma=8$. In their work, Miroudot et al. (2013) underline how the elasticity of substitution for the services market might be lower, since services are more 'ad hoc' than goods. Nevertheless, they use the same elasticity suited for the goods market aware that they might underestimate trade costs.

1.7 Sources of trade costs

The measure of trade costs is truly comprehensive in that it includes all costs occurred when participating in international trade exchanges. It consists of a wide range of direct and indirect components linked not only to transportation costs and tariffs, but also to all those costs that arise when trading goods internationally, like different currencies, languages, cultures, logistics services, imports and exports bottlenecks, etc. Anderson and van Wincoop (2004, p. 692-693) find that "*A rough estimate of the tax equivalent of "representative" trade costs for industrialized countries is 170 percent. This number breaks down as follows: 21 percent transportation costs, 44 percent border-related trade barriers, and 55 percent retail and wholesale distribution costs [...]. An extremely rough breakdown of the 44-percent number reported above is as follows: an 8-percent policy barrier, a 7percent language barrier, a 14-percent currency barrier[...], a 6-percent information cost barrier, and a 3-percent security barrier [...]*". They were the first to provide an approximation for trade costs basing the estimate on an extensive review of the literature on trade costs and of available information on trade costs.

Arvis et al. (2013b) separate the sources into two main categories. The first group of trade costs refers to all those bilateral factors that contribute to separating or to connecting trade partners and that are *exogenous* to countries' decisions or to policy choices. These factors are basically represented by the geographical distance between partners, by the time or costs required to transport goods and by all those common features that trading partners share (like language, culture, history, currency, border, being part of the same economic community or of the same trade union). The second set of sources of trade costs include all those costs that refer to the single exporting or importing partner and are *endogenous* to national choices and policy strategies. These costs measure the extent of the country's barriers and the complexities of its borders. They include tariff and nontariff[63] measures, all those costs connected with logistics performance (like costs, time and certainty of delivering the goods) and trade facilitation bottlenecks (such as the control at the frontier), and all those factors linked to the international connectivity and transportation (like maritime, air and land services). In view of

[63] "*Non-tariff barriers (NTBs) refer to the wide range of policy interventions other than border tariffs that affect trade of goods, services and factors of production. Most taxonomies of NTBs include market-specific trade and domestic policies affecting trade in that market. Extended taxonomies include macroeconomic policies affecting trade*" (Beghin in "The New Palgrave Dictionary of Economics", second edition, 2008).

following Chapter 4, it is possible to add one more category to the two trade costs sources identified by Arvis and co-authors. A more precise disaggregation might include all those factors that are specific to the exporting or importing partner but are exogenous to policy results or country decisions. This category might refer to all those geographical features linked to country, like percentage of rugged territory, having an access to the sea, being or not a landlocked country, being or not an island-state.

A different classification is between *fixed* and *variable* trade costs. *Fixed* trade costs are those costs that are paid once and are typically an investment, such as the investment to achieve a product standard or a regulation in the foreign market. *Variable* trade costs are paid for each shipped unit, like transportation costs (OECD, 2015).

Sources of trade costs and their extent may vary considerably between physical goods and services. In most countries tariffs are the most ordinary component when considering goods, but they account just for a small part of the total level of trade costs.[64] The main source is represented by other costs, mainly nontariff measures (such as product standards, domestic or international regulations, import quotas, exports restraints, etc.). As affirmed by Moïsé and Le Bris (2013), nontariff barriers (NTBs) have started to play an important role in international exchanges during last two decades with the increasing multilateral negotiations settled in international communities (such as European Union, World Trade Organisation, etc.). Institutions and the business environment also matter in the total amount of trade costs, because they make it more costly to do business abroad. For services, the sources and their importance in the computation of trade costs vary. Although merchandising goods dominate international trade, services represent 75 percent of domestic GDP in OECD countries and 20 percent of world trade.[65] They include transport services, logistics facilities, ICT, telecommunication, finance and insurance, business, legal and accounting services, tourism, construction and engineering, etc. For services, transport costs are basically null. The most significant element of trade in services is the domestic regulation and it may represent a friction if these laws and directions are not clear and predictable. They include all those rules and directives in land ownership, foundation of foreign companies, migration policies, and involve different sectors (banking, tourism, education, transport, etc.). Another important source of trade barriers for services are phone network or internet connectivity. Many firms pass through the web or widely employ phone or other telecommunication providers to deliver their services. For tourism, the existence of functioning transport infrastructure and

[64] OECD (2015).
[65] Moïsé and Le Bris (2013)

other facilities is essential to attract tourists; poor infrastructure system and lack of other services represent an important obstacle to this service.

Recent literature has attempted to decompose the measure of trade costs into its single sub-elements and to measure the relative impact of the different sources of trade costs using a regression approach.[66]

One of the first works providing a decomposition of the trade cost measure[67] into a range of sub-components is the paper of Chen and Novy (2011). They use data on 163 manufacturing industries in 11 European Union (EU) countries over the period 1999-2003 and try to explain the variation of trade barriers both across industries and across countries. As they assert, the main wedges of trade between countries are represented by the geographical distance and by more policy-related factors, such as the participation to common trade unions.[68] When looking at industries, the amount of Technical Barriers to Trade (TBTs)[69] represents the major issue for market integration across industries. Chen and Novy (2011) use a micro-founded measure of bilateral industry-specific trade frictions (the inverse of bilateral trade integration) as dependent variable and three sets of covariates: a set of geographical variables and transportation costs, a set of policy-related factors and a set of productivity variables and other costs (such as fixed costs to export). According to the authors, the main factors which make trade across nations costly are geographical impediments: trade frictions between two countries sharing a common border are 7 percent lower than between not contiguous countries; 14 percent lower if they share the same language. Being part of the same economic community contributes to making trade less expensive: trade barriers between two members are 10 percent lower than those between two non-members. For industries, regulations and standards constitute the chief obstacle to exporting in foreign markets: TBTs are too strong, creating obstacles to trade internationally; also high transportation costs contribute to making trade costly, especially for those industries producing heavy-weight products.

[66] A similar approach will be applied in Chapter 4 of this thesis.

[67] The trade cost measure employed has been obtained according to the top-down approach by Novy (2013).

[68] For the EU countries, the Schengen Treaty in the last three decades and all other agreements signed after the 1958, when the European Economic Community (CEE) has been established, have gradually removed trade barriers across countries, making trade in the EU more integrated. On this point, Chen and Novy (2011, p. 207) affirm that "*The case of the EU is appealing since trade integration is expected to be strong among its member states for two reasons. First, these countries have succeeded in dismantling many restrictions on trade, including tariffs and quotas that were completely eliminated by 1968. Second, the situation has been further reinforced by the implementation of the Single Market Programme (SMP), launched in the mid-1980s*".

[69] The Agreement on Technical Barriers to Trade has been negotiated in 1994 during the Uruguay Round, and, according to the agreement, TBTs aims "*[...] to ensure that technical regulations and standards, including packaging, marking and labelling requirements, and procedures for assessment of conformity with technical regulations and standards do not create unnecessary obstacles to international trade[...]*" (Source: https://www.wto.org/english/tratop_e/tbt_e/tbtagr_e.htm#Agreement, accessed 31 August 2016).

Arvis et al. (2013a) follow Chen and Novy (2011) extending the analysis to a larger number of countries, including other possible sources (air and maritime transport connectivity, logistics services, trade facilitation, and behind-the-border regulatory barriers). The contribution of Arvis et al. (2013a) is huge in terms of countries covered,[70] variables included and inference provided. They perform a cross-sectional regression using data for 2005 in order to maximise data availability, employing as dependent variable the bilateral trade cost between country pairs and employing fifteen covariates typical of the gravity model literature. They comprise both geographical and historical factors (distance, common border, common ethnographic language, common official language, common coloniser, colonial relationship, belonging to the same country, being landlocked) and traditional trade policies variables (tariffs, being members of a RTA,[71] exchange rate, maritime connectivity, air connectivity, logistics performance, entry costs in foreign markets). According to the authors, physical distance constitutes the major wedge for international exchanges between countries. Other geographical variables also have an effect on trade costs, but the distance remains the main determinant of overall trade costs. Trade policies also play an important role in trade costs: maritime transport connectivity and logistics performance have a considerable influence on making trade costly. Moreover, when combining them together, the extent of their joined effect is similar to that of geographical distance. Nontariff measures and other non-classical forms of trade policies also impact trade costs: nowadays nontariff barriers play a bigger role than tariffs in making trade more expensive, and the membership to trade community or RTA are helpful in cutting tariffs and reducing trade costs. These results apply to both manufacturing and agriculture sectors. The authors also assess how large is the contribution of each source to total trade costs by presenting standardised regression coefficients (betas[72]). Geographical separation and shipping connectivity have the biggest effect on trade costs: a one standard deviation increase in bilateral distance is associated with about a 0.45 standard deviation increase in trade costs for manufactured goods, 0.40 decrease when considering the maritime connectivity. The logistics performance generates a 0.20 standard deviation reduction and the other sub-components generate an effect of less than 0.15 standard deviation. Finally, Arvis et al. (2013a) attempt to measure the sources of trade costs

[70] 178 developed and developing countries.

[71] RTA stands for Regional Trade Agreement. The RTAs are reciprocal trade agreements between two or more partners which aim to increase economic integration and reduce barriers to trade, and have become prevalent since the early nineties (Source: https://www.wto.org/english/tratop_e/region_e/region_e.htm, accessed 31 August 2016).

[72] The beta coefficients are obtained dividing the parameters by the standard deviation of the related regressors (Arvis et al., 2013a).

identifying three trading groups: South-South, South-North, and North-North. North identifies all high-income countries and South includes all other countries. Therefore, the group North-North considers trade costs that incur in trade between developed countries, North-South takes into account trade costs between high-income and all other countries, and South-South looks at the trade costs between the non-high-income nations. The authors find that tariffs and other cost barriers affect trade costs in the South-South and in the South-North groups. Transport and logistics performance are important for both types of trade (manufacturing and agricultural) and for all groups, but the logistics seems to have a bigger effect on the North-North trade, highlighting how, in developed countries, efficiency, chain management and production networking are relevant. Foreign market entry costs are important for South-South and South-North trade relations, implying that trade policy reforms in developing countries are needed.

Novy (2013) performs a regression analysis for 13 OECD countries for a thirty years period (1970-2000), including among covariates two sets of trade cost proxies typical in the gravity model literature. The first group is composed by geographical variables (bilateral distance, sharing a common border, being an island), the second group contains institutional variables that capture historical and political traits (sharing common language, tariffs, being in a FTA,[73] being in a currency union). Using as dependent variable the indirect trade cost measure, Novy's estimates show that distance has an increasing effect on trade costs, whereas having a common border reduces trade costs. Being an island represents an advantage in terms of trade costs: island countries have lower trade costs and this can be explained, according to Novy (2013), with the fact that these countries, due their access to the sea, are more prone to trade internationally. Speaking the same language, thanks to the ease to conclude bilateral transactions and to the cultural closeness of people sharing a language, is associated with lower trade costs. Likewise, being in a free trade agreement and sharing a currency have a decreasing effect on trade costs, while tariffs, as expected, contribute to make trade more costly.

1.8 Concluding remarks

The chapter has provided a survey on trade costs, with a special focus on the indirect measure of trade costs computed according to the ‘top-down’ approach proposed by Novy (2013).

[73] FTA stands for Free Trade Agreement

The importance of trade costs in international economics lies essentially in the main role that trade costs play in global trade. Trade costs are a key factor in supporting or impeding trade across countries: high trade costs discourage firms from trading internationally and induce them to relocate their production domestically; on the other hand, low trade costs promote firms to export globally and to enter foreign markets. Moreover, trade costs are one of the principal reasons behind the waves of Globalisation and the decrease of trade in the interwar period, as well as the increasing trade in last decades. Last but not least, the still high value of trade costs, especially for developing countries, makes them one of the central issues of the political agenda of nations and of policy interventions.

During last fifteen years, the economic research has put much effort into examining the structure of trade costs, the methods providing the best trade costs estimation, the insight of inferring trade costs indirectly, the use of the best parameters to obtain finest approximations, the main sources behind the barriers to trade. This chapter has enlightened how, for almost half century, trade costs have been modelled as an ad valorem tax on the value of the good and, just in last two decades, researchers have begun wondering about an additive structure. The research on this issue has demonstrated how the choice between one of the two forms is impossible and both ad valorem and specific structure should be taken into account in theoretical and empirical works. Accordingly, the choice between a direct and an indirect method to compute trade costs is not feasible. On the one hand, the direct approach might produce a more accurate estimation on the extent of trade barriers, but the problems connected with data availability and with countries and time coverage make the work hard. On the other hand, the indirect approach does not suffer from limitedness of data, but the criticisms related to the use of an indirect method generate a more imprecise measure. These weaknesses behind the indirect method are overcome by the enormous advantages of the approach, which allow an extensive use and a wide application. The current wisdom of trade costs is enhanced, but further research is needed to explore new methods for obtaining more precise trade cost estimates, not only at the country level, but also at the regional one. Moreover, research should focus on the possible best actions to reduce barriers across countries. Geographical impediments are the main wedge for international exchanges: since is not possible to reduce physical distance, policy makers should find a key to make nations closer.

In conclusion, trade costs are noteworthy and they are the reasonable common explanation behind the key issues of international economics. They describe how countries' comparative advantage may vanish if trade costs are high. They explain how gains from exports may disappear if trade costs are high. They show how imports are no longer driven by preferences if trade costs are high. Finally, they are important because they help to answer the first and, maybe, the most significant question addressed by the theory of international trade: '*Why do countries trade?*'. And this is why trade costs are a crucial subject that asks for more research.

Chapter 2

The Persistence of History: A New Measure of Roman Roads

2.1 Introduction

Historians and economists have argued how historical events influence economic development and how they have been crucial for better institutions and government attitudes. History has been questioned in terms of persistence, and its long-term effect has been recognised for having important implications on actual economic patterns.

Starting from the work of North (1981), history has been found as having an important role in determining the current economic development, but it was during the first two thousands decade that several contributions gave new insights on historical variables as fundamental determinants of growth and current economic outcomes.

There are several channels through which history exhibits its persistent effect, but, as highlighted by Tabellini (2010), there is a widespread consensus for the legacy of history. Some economists argue that it is through institutions that history shapes the current economic performance. The institutions, beneficial for the economic development, are principally represented by limited government, incorrupt bureaucracy, legal system, low taxation and regulation. Economists came along with this conclusion centuries ago: Montesquieu and Smith, in the second half of eighteenth century, stressed the non-intervention of the Government, limited taxation and regulation as the best recipe for economic performance (La Porta et al., 1999).

Alternative to institutions, geography has been recognised by economists as not only a fundamental determinant of economic development, but also as having long-run effects on it. Recently, there has been a large debate whether geography has a direct persistent impact on growth or not. Bleaney and Dimico (2008) refer to a separation between '*pro-geography*' and '*pro-institutions*' economists. Authors like Knack and Keefer (1995) and Hall and Jones (1999) stress the importance of institutions on affecting growth in a persistent way, meaning that geography affects growth indirectly only through institutions; others, like Diamond (1997), Olsson and Hibbs (2005), Sachs (2003), find in geography a key and direct explanation for having long-run effects on economy.[1] The well-known work of Acemoglu et al. (2001) emphasises that only institutions matter for long-run economic growth, and that, once institutions are controlled for, geography has no significant direct effect on income. On the other hand, Sachs (2003) finds that geographical variables have a direct power in explaining the persistence in economic development.

Nunn (2009) refers to the debate in a more broad way, distinguishing between '*pro-geography*' and '*pro-history*' economists. The motivation behind the divergence lies in the fact that geography "*affects human actions in the past as well as today. In other words, in addition to affecting income directly, geography also influences history, which in turn affects current income*" (Nunn, 2009, p. 86).

Limão and Venables (2001) place their work between infrastructures and geography. They show the importance of infrastructures in determining transport costs and, consequently, bilateral trade, highlighting that remote, isolated and landlocked countries face higher disadvantages than coastal or island countries. A poor infrastructure system accounts 40 percent of predicted transports costs for coastal nations, 60 percent when considering landlocked countries. The key role of infrastructure for trade is strongly underlined by the work of Anderson and van Wincoop (2004). They highlight how poor infrastructures negatively affect trade and time costs. Their impact on trade costs influence in turn international trade volumes, and this impact differentiates across countries. Banerjee et al. (2012) emphasise that transportation infrastructures are considered one of the main determinants of growth and development. The reason for this is twofold. First, because to benefit from markets and from ideas it is necessary to reach them; second, because periods of development and economic growth in Western Europe, Japan and the U.S. were characterised by big infrastructure projects, like railroads. Aschauer (1990) describes the importance of infrastructures for the quality of life and for the economic performance, highlighting that numerous past infrastructure investments had a significant role and a large positive effect on

[1] For these authors geography works through several channels, like climate, disease environment, soil quality, access to markets, availability and productivity of labour and other factors of production (Engerman and Sokoloff, 2002).

people's life, improving their wealth in terms of health, free time and economic chance. The link between infrastructure projects and economic performance was particularly strong during the ‘golden age’: road-building projects and motorways considerably contributed to economic growth in the 1950s and 1960s.

The spirit of this chapter is twofold. On the one hand, it aims to stress the importance of history in influencing current development levels of countries or regions. On the other hand, it looks at a particular historical period: the Roman empire and its legacy in terms of infrastructures. As pointed out by Wahl (2015), historians and economists have proposed different channels through which the Roman domination might have left a mark on later developments. The performing institutions, the law and legal systems, the uncountable bishop residences, the urbanisation patterns, the developed market economy, the big infrastructure projects (bridges, aqueducts, canals, ports) and the extended road system represented the strong points of the Roman hegemony and were the basis of the survival of the empire for more than seven centuries. Although research on the Roman domination and its heritage on present development and income is significant, the empirical analyses on this subject have increased especially during last years. In this regard, the present chapter is completely devoted to the empirical application performed in Chapter 3, where the old Roman infrastructure is studied to assess its long-term effects on current trade costs of Italian provinces. The analysis is carried out using a novel measure of Roman roads specifically computed for the Italian provinces. The originality of this measure is that it can be easily implemented to all territorial levels[2] and quickly extended to advanced investigations.

This chapter consists of nine paragraphs. The first four paragraphs debate about the persistent effect of history, giving account of a new stream of literature that deals with the link between historical factors and present-day economic outcomes, referring to those channels (institutions, infrastructures, geography) through which history exhibits its impact over the centuries. The literature has been divided into four parts. Paragraph 2.2 provides a review of the traditional contributions on the long-terms effects of history, including as well some novel recent works and mentioning those papers that, on the contrary, refer to historical events which did not produce long-lasting effects today. Paragraph 2.4 looks at Italy and collects the contributions on the persistent effect of history on the sole Italian territory, disregarding, since the focus in following Paragraph 2.5, the glorious period that corresponds to the Roman domination. Paragraph 2.4 goes more into detail and considers that strand of literature which

[2] Countries, macro-regions, regions, provinces, municipalities.

retrieves in historical infrastructures a legacy in current development. Paragraph 2.5 is entirely devoted to those works that study the persistence of the Roman road system with a glance to that literature that evaluates the influences of the Roman domination other than the Roman roads. The constitution of a wide road network, considering the main features related to the development of the road network project and focusing on the state of mind behind the creation of an interconnected, efficient and cohesive empire, is examined in Paragraph 2.6. Paragraph 2.7 gives account of the important data source of the Roman road network represented by the McCormick et al. (2013) shape file used to compute the new measure of Roman roads. Paragraph 2.8 is the core section of this chapter: it describes the novel data sets produced and the newly created Roman road measure in kilometres by Italian province. A descriptive analysis is also presented. Some concluding remarks are provided in Paragraph 2.9. The chapter ends with Appendix 2.1, which presents the procedure and the technicalities behind the construction of the Roman road measure.

2.2 The persistent effect of history

In the last fifteen years a new strand of economic literature has grown. This literature aims to highlight how important is the role of history for the current economic development. The contributions flourished so rapidly that in 2009 Nunn produced a very complete and accurate survey on the persistence of history providing a precise distinction between the works that marked the origin of this literature, the more recent works and the channels through which history performs.

This paragraph distinguishes between those works that can be called, in the words of Nunn, the seminal contributions, some more recent articles that provide new insights to the growing literature and other contributions which do not retrieve in specific historical facts a persistent effect. This triple distinction has been reflected in three distinct subparagraphs. Subparagraph 2.2.1 provides a brief review of the six works that traced the birth of the literature on the legacy of history on economic outcomes. Subparagraph 2.2.2 presents three selected novel contributions that are, according to the author's perspective, particularly significant in terms of originality and linked to next empirical application in Chapter 3. Finally, Subparagraph 2.2.3 considers those works that, investigating on the persistence of specific historical facts, do not observe and find long-lasting effects of history on modern economic results. Table 2.1 summarises all these contributions, providing an overview of the main elements and aspects of each work.

Table 2.1 Seminal and more recent contributions to the literature on the persistent (and not) effect of history on present economic outcomes

Author(s)	Aim of the work(s)	Historical event	Main idea	Identification strategy	Quantitative method	Main finding(s)
Engerman and Sokoloff (1997, 2000, 2002)	To study the long-run economic growth and development across the Americas	European colonisation (from 16th century)	Factor endowment as long-term driver of institutional and economic development	-	Qualitative method. Historical and descriptive analysis	1. Differences in wealth, human capital, and political power rooted in the factor endowments 2. Early differences preserved by type of institutions
La Porta et al. (1997, 1998, 2008)	To establish whether the legal protection investor is a determinant of financial development	European colonisation (from 16th century)	British common law systems as long-term driver of institutional and economic development	Former American colonies legal origin as instrument for the protection of investors rights	OLS estimator for assessing the level of protection by law system. Two-stage procedure for evaluating the effect of legal protection on financial development	1.Legal rules protecting investors vary among legal origins: common law (rather civil law) countries more protective of outside investors 2. Legal tradition typically introduced by conquest or colonisation
Acemoglu et al. (2001, 2002)	To determine the impact of institutions on long-term economic performance	European colonisation (from 16th century)	Less deadly disease environment as long-term driver of performing institutions (via European settlements)	Mortality rates of first European settlers as instrument for current institutions (via early institutions)	Two-stage least square estimator for assessing the effect of institutions on performance (GDP)	1. Colonial experience as one of the many factors affecting institutions 2. Institutions of private property (rather extractive institutions) as long-run driver for economic growth
Davis and Weinstein (2002)	To examine the variation in economic activity across regions	Allied bombing of Japanese cities during World War II	It is possible to study the distribution of economic activity across space relying on the experience of one country	Deaths and destruction per capita (i.e. effects of bombing) as instruments for population wartime growth rates	Two-stage least square estimator for assessing the effect of wartime growth rates on growth in city population	1. Although the magnitude of destruction, most Japanese cities returned to pre-war populations 2. Distribution of city sizes seems to be highly robust to temporary shocks even of great magnitude
Dell (2010)	To use variation in the assignment of a historical institution in Peru to identify the channels through which institutions affect current economic outcomes	Mita (extensive forced mining labour system) in effect in Peru and Bolivia between 1573 and 1812	Land tenure and public goods as channels through which institutions produce persistent effects	Use of the historical mita boundaries given by distance and elevation to distinguish treatment and control samples	Regression Discontinuity (RD) approach for evaluating the long-term effects of the mita on household consumption and children stunted growth	1. Long-run negative effect of mita on household consumption 2. Long-run positive effect of mita on childhood stunting 3. Land tenure, public goods, market participation as channels through which institutions perform
Miguel and Roland (2011)	To investigate the effects of the U.S. bombing on lung-run economic development in Vietnam	Vietnam War and the U.S. involvement between 1965-1975	Long-run impacts if there are poverty traps. Poverty traps might be prevented by factor mobility within country or government redistribution of capital	Use of the distance from a district to the 17th parallel-border between North and South Vietnam as instrumental variable for bombing intensity	Two-stage least square estimator for assessing the effect of bombing on per capita consumption levels and growth, population density, physical and human capital investment levels	1. No robust long-run effect of U.S. bombing on local poverty rates, consumption levels, or population density 2. Vietnam's rapid recovery from U.S. bombing suggests the implausibility of a poverty trap
Redding et al. (2011)	To provide empirical evidence that industry location is not uniquely determined by economic fundamentals	Division of Germany after World War II and the reunification of East and West Germany in 1990	There exist multiple steady states in industry location	Use of the division of Germany and focus on an air hub as natural experiment	Difference-in-differences approach to show that the treatment effect of division on the location of the hub is statistically significant and the treatment effect of reunification is no statistically significant.	1.The relocation of West Germany's main airport hub from Berlin to Frankfurt after WWII division persisted after reunification in 1990 2. Multiple steady state locations 3. Differences in fundamentals dominated by the sunk costs of building the hub
Valencia Caicedo (2014)	To explore the path dependence in economic development of other (not English or French) colonial arrangements and institutions	Establishment of Guarani Jesuit Missions in South America (1609-1767)	Places closer to missionary districts have higher current income thanks to these institutions	Computation of Altonji ratios, use of abandoned missions, Franciscan Missions as comparison to control for the endogeneity of missions location	OLS estimator with fixed effects for assessing the effect of the Jesuit Missions on contemporary outcomes	1. Long-lasting positive effects of Jesuit Missions on income and education 2. Human capital as the main channel of transmission 3.Cultural mechanisms can sustain differences
Cantoni (2015)	To assess the effects of Protestantism on long-run economic growth	Holy Roman Empire between 1300-1900	Protestantism has a positive effect on economic growth	Holy Roman empire as natural experiment. Further strategy: distance to Wittenberg, where Martin Luther presented his 95 theses, as an instrument for Reformation	Difference-in-differences approach to capture differentials in city growth between Protestant and Catholic cities	1. No evidence of any differences in the long-run performance of Protestant and Catholic regions 2. No evidence of long-run effects of Protestantism on economic growth
De Benedicits and Pinna (2015)	To evaluate the geographical condition of insularity and the role of connectedness in influencing islands' trade costs	British, Dutch, French and Spanish colonisation between 1750-1850	Historical interactions shaped a culture of openness (institutional connectedness) able to reduce the costs of being an island	More than 5000 trips between 1750-1850 to measure how connectedness has built in time: islands most likely territories encountered by ships navigating the sea.	OLS and Hausman-Taylor IV estimators for assessing the effect of insularity and geography on trade costs	1. Spatial connectedness: for islands their position with respect to their possible trade partners is crucial 2. History shaped the geographical distance and connectedness
Waldinger (2016)	To examine the effects of physical and human capital for scientific productivity	Nazi Germany and World War II	Both human and physical capital are important in determining productivity	Use of the dismissal of Jewish scientists in Nazi Germany as human capital shock and Allied bombings during WWII as physical capital shock	OLS estimator with fixed effect and interactions variables for assessing the effect of dismissal and bombing on scientific productivity (publications)	1. Short-run: both human and physical capital shocks have a negative effect on scientific output 2. Long-run: negative effects persist in the long-term only for human capital
Saing and Kazianga (2017)	To investigate long-term economic effects of armed conflicts	The U.S. bombing in Cambodia between 1969-1973	Armed conflict may cause both short-run and long-term consequences	Use of bomb intensity to measure the destruction of conflicts	Difference-in-differences approach for assessing the impact on different outcomes at the micro level (households) of bombing	1. U.S. bombing reduced years of schooling , increased total number of births and age at first marriage among female 2. Any long-run effect on earnings, employment and women's height

Source: Author's elaboration

Some more few words are needed to explain the reasoning behind a similar categorisation. The inclusion of the seminal contributions has the precise aim, on the one hand, to list what are the benchmarks of the literature; on the other hand, they represent a good starting point in terms of identification strategies and estimation techniques for the empirical application of following Chapter 3. The extension of the literature review on more recent works has, instead, the purpose to highlight, first, how the focus on historical events, which triggered long-lasting effects on current economic outcomes, changed across years and how a big variety of works grew up following different identification paths. On these bases, the choice of these novel contributions has been not trivial: those works more related to the scope of this thesis and published or in progress after the survey of Nunn have been preferred. Lastly, a brief focus on those historical events that did not affect current economic outcomes helps to complete the literature overview.

2.2.1 The six pioneer contributions

According to Nunn (2009), six works and nine authors marked the birth of the literature on the persistence of history on present economic results: Engerman and Sokoloff (1997, 2002), La Porta, Lopez-de-Silanes, Shleifer and Vishny (1997, 1998), Acemoglu, Johnson and Robinson (2001, 2002). These six works share three main points which represent the key elements of their analyses as well. First, they explain differences in productivity and per capita GDP as long-term consequences of history and better/worst historical institutions, via existing institutions. Second, they are the first to draw attention to a historical event as main explanatory factor of their analysis looking at the same historical period: the European colonisation from the sixteenth century. Third, all of them retrieve in the quality of institutions a long-term determinant of economic development and argue that historical events (like the colonial rule) have a big role in shaping institutions. What distinguishes them is the different source of colonisation which shaped institutions they claim. While Engerman and Sokoloff suggest that the origin of the missed development of institutions that promote long-term economic growth should be retrieved in different endowments of land and geography, for La Porta et al. the cause lies in the differences between legal systems based on British common law and those based on Roman civil law, whereas for Acemoglu et al. what generated the settlement of European colonisers, which in turn encouraged growth-supporting institutions, was the subsistence of a less deadly disease environment.

The case of the Americas is quite attractive and explains why the earliest studies focused on them. In fact, after the arrival of the Europeans, there were not large differences across the economies in the New World and the absence of disparities lasted for more than 250 years. During almost all eighteenth century the prosperity of the Caribbean and South America was in line with the wealth of North America (the U.S. and Canada); for some colonies, like Barbados or Cuba, per capita incomes have been estimated as 50 and 67 percent higher, respectively, than those of the U.S. (Engerman and Sokoloff, 2000). It was only in the late eighteenth-early nineteenth century, with the industrialisation process, that the disparity between North and Centre-South America appeared: the U.S. and Canada began to industrialise earlier than the southern colonies and this allowed a major growth over the long run. North (1987) was one of the first ones to put his attention on the differential paths of development between North and South America, suggesting that the U.S. and Canada had experienced a long-run development thanks to the British institutions, more able to promote sustainable growth than the Spanish or other European institutions. The superiority of the English institutional heritage and the match between Protestant principles with market institutions were only one of the reasons used to explain these differentials in growth rates among the American colonies; other authors pointed out on cultural and religious factors (Engerman and Sokoloff, 2002). Although these 'primordial' interests on the diverse development of America, they did not mark the birth of the literature on the persistence of history.

Differently from the earliest explanations, the studies of Engerman and Sokoloff (1997, 2002) find in what they call 'initial conditions' the cause of the diverse development paths of the American colonies. This primary difference in factor endowments, in land and geography required the use of slaves for the cultivation of sugar, cotton and other crops encouraging the constitution of a society characterised by economic and political inequalities: a restricted and privileged group of European descendent owned a vast share of the total wealth, human capital and political power. The English and French colonies in the northern part of North America were unfamiliar with this kind of organisation, supporting instead all-inclusive societies.

The centrality of Engerman and Sokoloff's contributions in providing a specific and detailed explanation on the source that generated the constitution of diverse performing institutions acquires significance in the view of next Chapter 3. To support their key idea, Engerman and Sokoloff perform a deep historical research in order to accurately understand the basis of the different starting endowments able to explain the dissimilar economic development across the Americas. Accordingly, the empirical application on Roman roads extensively exploits historical sources to better address the central subject of the investigation and understand what are the mechanisms subtending the historical event considered.

For La Porta et al. (1997, 1998) the differences in development and the economic and political inequalities originate from what they call 'legal origins': the historical origin of a country's laws is strictly connected with its legal rules and regulations, as well as with economic outcomes. In other words, the legal origins can be considered as 'persistent systems of social control of economic life' having effect on the legal and regulatory structure of the society and on economic outcomes. They support this idea explaining the differences between the British common law and Roman civil law legal systems: in countries where common law legal systems are in force, outside financial investors benefit from more protection than in nations shaped by the civil law system, and the evidence show that legal investor protection is a good predictor of financial development. La Porta et al. find the legitimacy of this idea in the colonisation of the American continent. The northern part of the Americas, basically those territories composing nowadays the U.S. and Canada, were under the British or the French control; whereas Portugal and Spain dominated in the southern part of the continent. While the U.K. is a typical common law legal system, France, Spain and Portugal are civil law systems; hence, the British colonies were patterned on the common law system, while the French, Spanish and Portuguese on the civil one. The establishment of common and civil law in the colonies was not insignificant: colonies were shaped not only by the system of rules, but human capital, legal ideologies and the economic system were affected too. In other words, the transplantation of the legal system entailed the creation of a society strongly based on the legal system of the former coloniser. What is important to underline is that the works of La Porta et al. and their 'legal origin theory' do not claim the superiority of common law system on the civil law ones, but they find in the common law system a better system able to promote sustainable growth in the long-run and to explain why the northern territories of America have generated an extended development across the centuries so that to become the more productive economies today.

As La Porta et al. (1997), Acemoglu et al. (2001, 2002) confirm that differences in institutions are at the root of different development across countries. While La Porta et al. retrieve in common and civil law systems the source of these disparities, Acemoglu et al. explain the difference in institutions as the consequence of the feasibility in the European colonisation. The settlement of the Europeans in the new world was basically determined by the survival conditions of the colonisers in those territories. In colonies where Europeans could survive, were established institutions able to promote extended growth and development and to guarantee the rule of the law and the respect of the property rights. Where, instead, the settlement was not feasible due the more deadly disease environment, colonisers were not encouraged to set up institutions enforcing the law and promoting investments. In other words, Europeans employed different colonisation strategies and established different institutions according to the quality of environment they found. The empirical analysis provided by Acemoglu et al. confirms that America territories with high European mortality rates were weakly touched by the colonisation process. In these territories the European settlements were little and institutions were not developed. Moreover, they find high correlation between past and current institutions and a large effect of institutions on per capita income. In the view of Acemoglu et al., institutions are like a ‘black box’ and the colonial experience is only one of the factors affecting their functioning.

The work of Acemoglu, Johnson and Robinson of 2001 is particularly significant in the economic literature for the quantitative method they use. They rely on the instrumental variables approach to deal with the possible endogeneity problem of institutions. On these bases, they use the mortality rates of the European settlers as instrument for current institutions in America. The reasoning behind this is that less deadly disease environments (with lower mortality rates) allowed the Europeans' settlement and the consequently establishment of early institutions. Early institutions are strongly correlated with current institution and the latter impact on the present income. Chapter 3 exploits a similar approach for the effect of the Roman road infrastructure on current trade costs.

2.2.2 Three novel contributions

During last fifteen years, new articles contributed to giving innovative insights on the channels through which history works. A flourishing literature started to grow taking into account the impact of various and original historical shocks. A complete and comprehensive review of all these works is impossible; here, three of these original contributions have been selected and are presented, starting from the latest one and concluding with to the ‘oldest’ one. The choice of these contributions has been driven by the innovative character, the originality of the historical episode analysed and by the connections that these works have with following Chapter 3.

The first contribution is the novel work of De Benedictis and Pinna (2015). They look at the historical trade routes and put their work in the midpoint of the debate between ‘geography matters’ and ‘history matters’. The double ‘geographical-historical’ dimension of the paper can be retrieved in the data they use. One the one hand, they classify countries according to their ‘degree of insularity’, distinguishing landlocked countries from island-states and comprising three additional intermediate levels of insularity: coastal countries with islands, coastal countries without islands and coastal countries with islands representing a negligible part of the total national territory. According to De Benedictis and Pinna's measures, being an island is characterised by costs and is linked to higher trade costs; these costs reduce with connectedness and with what they call ‘institutional connectedness’, i.e. the development of a culture of openness due to repeated historical interactions with merchants from mainland. On the other hand, they use British, Dutch, French and Spanish ships logbook records for the period 1750 to 1850, describing the major navigation routes in the eighteenth and nineteenth centuries, in order to identify the islands touched by the routes, the islands representing anchorage points for the vessels or the islands used as harbour stops by the ships. De Benedicits and Pinna show how geographical features determined participation to historical trade routes between 1750 and 1850, shaping openness and connectedness. Using islands as a natural experiment in order to evaluate the influence on trade costs from geography and historical events, they prove that, once geographical features (and physical distance) have been taken as given, the other factor that contributes to building connectedness can be retrieved in history. The work of De Benedicits and Pinna is particularly important in the view of the empirical investigation on Roman roads in following Chapter 3, since both contributions exploit the indirect measure of trade costs computed according to the ‘top-down’ approach proposed by Novy (2013) as main dependent variable of the analysis. De Benedictis and Pinna use the classic country level view with a particular focus on island countries, while Chapter 3 exploits the sub-country (provincial) level perspective.

Another interesting and recent contribution is the work of Valencia Caicedo (2014), who looks at the Jesuite Missions of the seventeenth and eighteenth centuries in South America as a reasonable factor affected by persistence and able to explain the current income and education level in Argentina, Brazil and Paraguay. The Jesuite Missions, according to Valencia Caicedo, represent an ideal case study to examine the persistent effect on income, human capital and culture. In fact, due political reasons, Europeans decided in 1767 to expel the missions from the southern colonies impeding a continuative effect in the following centuries. Differently from other missions (such as the Franciscan Missions), the Jesuite Missions had an important role in terms of education externalities and instruction for children, created human capital spillovers and generated skilled labour training for the adults; moreover, they created persistent occupation flows. In other terms, the Jesuite Missions offered something more than the mere religious scope: they generated education and provided technical training in their conversion. In fact, Valencia Caicedo finds that in territories closer to those that were missionary districts the level of current income, education and literacy are higher, people work more hours and specialise in manufacturing or services rather than agriculture. The work of Valencia Caicedo is the first one to provide a study on the long-run economic consequences of the Jesuite institutions; even though historians and philosophers argued about the positive effect in terms of industriousness and education, the knowledge about the effects of old religious organisation in the modern economy is still little. Interesting point of the Valencia Caicedo's contribution is the importance retrieved in the cultural mechanisms in sustaining persistent outcomes. In his view, culture is the device through which differences in income and human capital persist over time and, more in general, is the mechanism through which history performs its persistent effect. These thoughts are in line with the work of Nunn of 2012 and integrated in following Chapter 3 as well.

A third significant contribution in the literature of the persistence of history is the work of Dell (2010). Dell studies the long-term effects of the mita on actual economic incomes. The mita was a forced mining labour system in effect in Peru and Bolivia between 1573 and 1812. The Spanish established the mita during the colonisation of Latin America and required that one-seventh of adult male individuals of over 200 indigenous communities were sent to work in silver and mercury mines. Making use of the fact that the borders of the mita area experienced a variation, Dell compares outcomes of mita and non-mita districts close to the boundary in order to estimate the effects caused by the mita in the long-run. The study shows that the mita system negatively affected current economic development, household consumption and stunting in children. Territories that were part of the old mita districts have nowadays 32 percent lower household consumptions that non-mita districts. Moreover, the forced labour system produced a negative effect in terms of education, road networks and connectedness terms. During the mita, the long-lasting presence of big landowners in non-mita districts generated a stable land tenure system that encouraged the provision of public goods, like roads, thanks also to the landowners' ability and contacts with lobbies to secure roads. On the other hand, in the mita districts a similar provision was not favourable and this may explain why actual mita districts remain less integrated into road networks. Dell's contribution in the view of Chapter 3 is interesting for the strategy she uses to identify areas that were subject to the mita and areas that were not affected. Geographical features, given by distance and elevation, marked the boundaries of the mita conscripts during sixteenth and seventeenth centuries. Dell exploits the similar features to distinguish the current treatment area from the control area. The permanent mark of geography is also used in Chapter 3 to deal with a possible endogeneity problem that typically has to do with the old and the modern infrastructure.

2.2.3 The missing persistent character of history

The review until now presented has illustrated the seminal and some more recent contributions of that strand of economic literature that argues that history is able to produce long-lasting effects, affecting past and modern economic outcomes. In more than fifteen years, a wide range of historical episodes has been investigated and significant persistent effects have been retrieved in their shocks. However, the mark of persistence, allocated to history, does not emerge in each context and for each event in history. Inside the economic research on the legacy of historical facts, some contributions bring to light those cases and circumstances in which history has not produced inheritance on present economy. These are episodes that generated impacts in the past, but are not affected by the 'lasting feature' able to cause an effect today. The remaining of this paragraph is focused on these contributions.

One of the first contributions, that, studying the impact of history today, does not find a long-run effect, is the work of Davis and Weinstein (2002). In order to investigate on the distribution of economic activity within a country, according to three chief theories able to explain the spatial pattern and variation across regions, they look at the distribution of the regional Japanese population in a time span of 8,000 years, from the Stone Age to 1998, and exploit a historical shock to assess whether temporary shocks have persistent effects on urban areas. The historical fact they use is the bombing of Japanese cities by the Allied forces during World War II. The empirical evidence coming from their instrumental variables approach, where deaths and destruction are used as instruments for population growth rates, highlights how big temporary shocks do not affect in a permanent way the spatial structure of the economic activity and how Japanese cities recovered their former populations in a short time. In fact, although the magnitude of devastation and of bombing during World War II, there is no proof of long-run impact on the relative size of cities: the urban areas were restored in a period of fifteen years.

The work of Miguel and Roland of 2011 shares with the paper of Davis and Weinstein the war subject. However, whereas for Davis and Weinstein the allied bombing of Japan during World War II is a tool to examine the spatial distribution of economic activity, in Miguel and Roland (2011) the historical issue is the core of their study. They are interested in investigating what are the long-term effects of the U.S. bombing on the economic growth of Vietnam. They start from a twofold perspective of theory, according to which, on the one hand, the destruction of the capital stock during wartime may lead to a poverty trap with a subsequent long-run underdevelopment, whereas, on the other hand, the economy is able to recover pre-war physical and human capital accumulation levels without long-term impacts on the economy. On these bases, they use the distance from a district to the 17^{th} parallel, the North-South Vietnam boundary, where, between 1965 and 1975, the U.S. army focused its bomb attacks, as an instrument for the bombing intensity. Relying on this identification strategy, Miguel and Roland find no empirical proof of long-lasting impacts on modern poverty levels, household consumptions, literacy rates, infrastructural standards and population density. Vietnamese districts recovered fast from wartime bombing, suggesting that no poverty trap occurred and, consequently, the doubtfulness of the poverty trap theory for Vietnam. Vietnam, such as Japan and Germany, benefited from performing institutions able to arrange reconstruction and wealth redistribution, which allowed to successfully overcome the drawbacks of war.

In line with the work of Miguel and Roland (2011), the very recent contribution of Saing and Kazianga (2017) looks at the U.S. bombing in Cambodia between 1969-1973 to explore the long-run effects of armed conflicts on development. Exploiting bomb intensity as a measure of war destruction and adopting a difference-in-differences approach to calculate the effects of bombing on a wide set of economic outcomes, they find that war negatively affected years of education, and increased the number of birth and the age at first marriage for females. However, there is no evidence of long-term effects on earnings, employment rate and women's height, corroborating Miguel and Roland's findings.

In 2011, Redding, Sturm and Wolf look at the division of Germany after World War II and the reunification of East and West Germany in 1990 to empirically explore the industrial location. In particular, the work of Redding et al. (2011) exploits a particular economic activity, the air hub, and a particular historical fact, the relocation of the main German airport from Berlin to Frankfurt after World War II, to study the pattern of the distribution of the economic activity. Two different classes of theoretical models predict the distribution of industrial activities. The first class argues that initial conditions, historical shocks, agents' expectations determine multiple steady state distributions. According to the second class, instead, institutions and endowments are the main factors affecting a unique steady state industrial location. The works of Davis and Weinstein (2002) and of Miguel and Roland (2011) above presented are two empirical proofs of the existence of a unique stable steady state, since the complete recover of pre-war cities size and wealth, suggesting a central role of economic fundamentals in determining industrial locations. However, they do not provide a proof against the existence of multiple equilibriums. The evidence emerging from the paper of Redding et al. (2011) is original. Redding and co-authors, starting from the missing restore of the leading air hub from Frankfurt to Berlin in 1990, after reunification of West and East Germany, find how the treatment effect of the division is statistically significant and how, instead, the treatment effect of reunification in 1990 is not statistically significant, suggesting a persistent character in the location of the economic activity after the movement. Redding et al. explain this result as a consequence of high sunk costs of building air infrastructures and argue that multiple steady states for airport locations are possible.

A further evidence on the absence of long-lasting impacts of history has been highlighted in previous Subparagraph 2.2.2, when exposing the work of Valencia Caicedo (2014). Valencia Caicedo argues that the foundation of Jesuite Missions during the seventeenth and eighteenth centuries is able to explain the current income and education level in South America, but this is not the case when considering the Franciscan Missions. Although Franciscan Missions have been established before the Jesuite ones and had the opportunity to locate in better sites, there is no evidence of long-lasting effects on modern incomes or literacy. Valencia Caicedo describes how these differences can be explained by considering the duty of the missions. Franciscans were more focused on reducing poverty and inequality. The Jesuits, instead, were more oriented on the technical and human capital formation. Moreover, the Franciscans were more subject to the Spanish colonial rule. The Jesuits, instead, were more able to obtain lower taxes and labour tributes from the Spanish government.

Another recent work is the one of Cantoni. Cantoni (2015) puts his work in that branch of the literature according to which Protestant, rather than Catholic, religions positively affect economic development in the long-run. On these bases, since the heterogeneity in religious beliefs in Germany during the Middle Age, Cantoni looks at the Holy Roman Empire between 1300 and 1900 as a natural experiment to empirically assess the relationship between Protestantism and economic growth. Using a difference-in-differences setup in order to measure disparities in urban growth between Protestant and Catholic cities and exploiting an additional instrumental variables approach, where the distance to Wittenberg, the Martin Luther's city, is used as an instrument for Reformation across Germany, Cantoni finds no long-run differences in economic performance between Protestant and Catholic cities. The explanation behind this original result might be connected to the works of Doepke and Zilibotti (2005, 2008), according to which the economic circumstances, rather than the religious beliefs, are the main drivers of individuals' personal improvement in terms of prudence, literacy and parsimony.

To conclude this review, an interesting and very recent contribution of Waldinger is presented. In the work of Waldinger (2016) the historical event is used to examine what is the effect of human and physical capital on scientific productivity. He looks at Germany under the Nazi rule and during World War II. Since the endogeneity problem linked to both inputs (most productive departments attract better scientists and invest in physical capital, and good researchers and advanced physical resources enhance productivity), two historical shocks are used to exogenously identify human and physical capital levels: the dismissal of Jewish scientists and the allied bombing. Waldinger, exploiting data on publications and citations for different years until 1980, finds that the layoff of many Jewish researchers (human capital shock) negatively affected both in the short- and in the long-run the knowledge productivity in departments and universities. The allied bombing (physical capital shock), instead, had a negative effect only in the short-run; persistent impacts are not retrieved in recent productivity. The main conclusion coming from the work of Waldinger (2016) is that both inputs are determinant in the production of scientific knowledge; however, whereas physical capital is able to restore in short time its pre-existing levels, negative shocks on human capital persist much longer.

2.3 The persistent effect of history in Italy

The focus on Italy and on the long-lasting effect of the Roman road network in following Chapter 3, highlights the need for a further investigation about which past historical episodes and facts impacted the modern Italian economy. Italian history is one of the richest in terms of confluence of events and of variety of influences and effects. Given its central and strategic position in the Mediterranean Sea and because of its northern union with the rest of the continent, Italy has represented for centuries the crossing-point to Europe, Africa and Middle East and was the place where a mixture of dominations, cultures and peoples converged.

Disregarding the legacy coming from the Roman domination in Italy, since the focus on next Paragraph 2.5, this paragraph aims to create a brief collection of the works that find in the Italian history a mark of persistence and a significant role in affecting modern economic outcomes.

The first work selected to open this review is the significant contribution of Robert Putnam with Leonardi and Nanetti of 1993, a long-term investigation about the development of institutions and their adaptation to the social context in the Italian regions. The work of Putnam et al. (1993) represents the seminal contribution in the study of the legacy left by history in Italy. Considering all facets of institutions (performance, development, the links with the past civic tradition and with the social capital), the Italian regional framework is, according to Putnam and co-authors, a perfect experiment to examine what makes institutions robust, receptive and representative. In their view, institutions have a double character: an active role in determining individuals' identities and strategies, and a dependent character because of the pervasive effect of history in shaping institutional performance. They point out, how the differences among regional governments in Italy are deeply rooted in history and how their current borders correspond to the boundaries of the historical dominations that ruled in the Italian peninsula during the Middle Ages. Indeed, the regional governments were established in Italy only in 1970; for more than a hundred years (after the Italian unification in 1861) the power was centralised. However, for a millennium and a half (after the dismissal of the Roman empire, but before the unification of Italy) the Italian territory was marked by several regimes and dominations that succeeded along the centuries and that were diverse between northern and southern regions. These dominations were responsible for the institutions they established in terms of social, cultural and political environments and arrangements. In this framework, Putnam and co-authors carry out a structured research focused on addressing the effects of institutional change. They find that the Medieval history, and the civic traditions and social solidarity that originated during this period, are decisive in explaining current Italian regional differences and in providing a further explanation, among the wide assortment of existing ideas, for the Italian dualism between the economically developed regions in the North-Centre and the less developed regions in the South. History and social environment are determinant factors for the success of institutions.

The remarkable insights emerging from the contribution of Putnam, Leonardi and Nanetti represent the starting point and the first background of all those works that, particularly in last years, started to show interest in the persistent effects of Italian history. Indeed, recently, a specific literature has grown exponentially and new appealing empirical works have started to provide new evidence on the fact that history is affected by persistence and performs its effects on modern Italian economic outcomes.

Starting from Putnam's conjecture about how current Italian regional differences in trust and cooperation originate from the historical independence that cities experienced during the Middle Ages, Guiso, Sapienza and Zingales (2016) are interested in understanding the source of historical shocks to institutions and, more in detail, whether these shocks perform a persistent effect directly (via formal institutions) or indirectly (via the effect they produce on mentality and culture). On this aim, they look at the Italian context and examine whether the differences in current civic capital between Northern and Southern Italy derive from the medieval independent city-states. Exploiting different measures on civic capital and identifying free Northern city-states using indicators of whether they were independent (e.g. free city in 1176, duration of independence, etc.), Guiso and co-authors find that the existence of independent city-states during the Middle-Ages positively affects civic capital today. The persistent character performs not only with the mere establishment of city-states, but their duration also has an important role, suggesting that, the longer the existence, the greater the effect. Since the constitution of the Norman kingdom in the South has not favoured the establishment of city-states, Guiso and co-authors are forced to compare current levels of civic capital only across Northern cities. In order to provide a counterfactual analysis, they rely on instrumental variables and on the orthogonality of two instruments (presence of a bishop in the city and Etruscan origin of the city) to prove that they do not affect cooperation in southern territories, corroborating the evidence that Northern cities, because of the independence during the Middle Ages, have today higher levels of civic capital.

The case of Italian cities and the role played by both institutions and geography in determining the long-run urban development is the main theme of the work of Percoco (2009). Analogously to Guiso et al. (2016), Percoco looks at the Middle Ages in Italy and, in particular, at the long period between 1300 and 1861, but, differently from the work of Guiso et al. (2016) that supports the *pro-institutions* view, the contribution of Percoco (2009) embraces both institutional and geographical perspectives in what he calls *institutions and geography* view. According to Percoco, Italy represents an ideal context to test both long-run effects of historical events and adverse geography on urban development, because of the high variability in terms of institutions, geographical characteristics and economic performance of cities. Exploiting data on city population, geography and quality of institutions for almost 600 Italian cities for a time span of almost 600 years, Percoco proves that both geographical (like accessibility, having seaports) and historical features (like the presence of universities) are important in shaping cities and in determining modern urban development.

Analogously to Guiso et al. (2016) and to Percoco (2009), Di Liberto and Sideri (2015) consider the long period of the Middle Ages in Italy and the institutions perspective. Di Liberto and Sideri, as it will be extensively examined in Chapter 3, are interested in examine the relationship between economic performance and the quality of government institution. They rely on the variability among the dominations that ruled in Italy during seven centuries (before the Italian unification) to construct two different instruments able to capture the exogenous variation within the Italian provincial institutional quality. Identifying for each province which sovereignty ruled during the hundreds of years period of the Spanish domination in Italy (first instrument) and measuring for each province how many years lasted each domination in the period between 1100 and 1700 (second instrument), Di Liberto and Sideri find that historical institutions are important in explaining differences in modern institutions and that the variability among the quality of institutions plays a significant role in determining economic performance of Italian provinces.

If the works discussed above are mainly focused on institutions, as the channel through which history performs its effect today, and if they all consider the period of the Middle Ages, as the epoch that largely contributes to explaining current economic patterns, the contribution of Pascali (2016) is quite innovative in both historical channel and time it looks at. Considering a specific period of the lengthy medieval history, the Italian Renaissance, Pascali documents and explores a pioneering channel: the creation and the development of local banks. The work of Pascali aims to test the conjecture according to which the Jewish community in the early Renaissance played an important role in the development of local banking, and this positive effect lasted until today, fostering current economic performance of a large number of Italian cities. According to Pascali (2016), Jews arrived in Italy during the Roman Empire, but, due to commercial reasons, they were expelled from Rome and deported as slaves in the rest of the Italian territory. If for about 1400 years (0-1400 A.D.) Jews were mainly small merchants, artisans or farmers, during the fourteenth century, because of the ban of lending money to earn a profit by the Catholic Church, Jewish people became moneylenders. Exploiting the Jewish population in 1500 as an instrument for the contemporary local bank (Jewish demography during the Renaissance is correlated with the local credit market in 1500, which, in turn, is correlated with current local banking, but not with contemporary per capita income), Pascali provides a significant empirical evidence. He documents that the Jewish presence in Italy during the Renaissance fostered the development of a financial credit system that lasted until today and that, in turn, generated a higher current income. In those areas where the current availability of credit system is larger, banking institutions promote firm productivity. Moreover, Pascali provides an explanation for the North-South Italian dualism: the expulsion of the Jewish communities from the Spanish territories (mainly in Southern Italy) has dampened the economic development of those areas.

This brief review on the persistent effect of history on modern Italian economy, ends with the contribution of Fontana, Nannicini and Tabellini of 2016. The three authors discuss the persistence of two important historical facts in Italy on political attitudes of citizens. In light of the fact that democracies that originate from civil wars inherit divisions and conflicts, affecting the functioning of modern political organisations, Fontana and co-authors look at the civil war and at the Nazi occupation, occurred in Italy in the last two years of World War II, to explore the effects on the new-born domestic political system. They find that, the stronger and the longer the civil war and the foreign occupation were, the higher the extreme left votes fraction was. This effect was not limited to the years just after the war, but it persisted until the early Nineties, with the end of the First Republic. According to Fontana and co-authors, there are two possible explanations behind these effects. On the one hand, the radical left, due the strong involvement in the fight against the Nazi forces during the wartime, has benefited from the voters' gratitude and appreciation. On the other hand, the communists had the opportunity to create and to arrange an active political organisation in those territories where the resistance movement lasted for a longer time.

2.4 The persistent effect of historical infrastructures

Recent literature has revealed interest in the effect of historical great transportation infrastructure projects on reducing trade costs, on playing a positive impact on productivity and on increasing the level of real income in trading regions.

Public investments in transportation infrastructure and their effects on income, exports and in education have been largely debated since the earliest projects in ancient times. Some politicians, historians or simply observers were sceptical on the positive impact that new roads and railroads might perform on development and wealth, others argued that these investments could have promoted beneficial effects on several spheres of the economy. Today, the positive consequences deriving from public investments in new transportation infrastructure are highly recognised, although policy makers claim for more proofs on the extent of their effects. The experience of historical big transportation constructions represents a helpful and practical evidence on what have been the favourable results of new transportation infrastructures.

During last five years, a growing number of contributions about the old historical infrastructure emerged. Starting from the work of Aschauer (1990), which pioneered the empirical evaluation of the effects of transportation projects, a younger literature on a variety of transportation infrastructures in different areas of the world has been developed: from the Inca roads in Peru to the railroads in Sweden, from the steamships in Argentina to the railroads in America, from the Chinese railroads to the railroad network in America and India. Table 2.2 gives a schematic overview of the contributions here examined.

Table 2.2 Contributions to the literature on the persistent effect of historical infrastructures on present economic outcomes

Author(s)	Aim of the work	Historical transportation infrastructure	Main idea	Identification strategy	Quantitative method	Main finding(s)
Banerjee et al. (2012)	To evaluate the long-run effects on wealth and economic growth of better access to transportation infrastructure	Railroads in China: constructed in late 19th-early 20th century	A greater access to transportation network has a long-run effect on trade, prices, localisation of economic activity and pattern of cities	Straight lines connecting historical cities and/or "Treaty Ports" and capturing old railroads as an instrument for the current railroads and coastal routes	OLS estimator for assessing the effect of railroads (via distance to the straight line) on GDP per capita, GDP growth, firm location and profits, household income	1. Being close to historical transportation infrastructure positively affects GDP per capita, income inequality, number of firms and their profits. No effect on income growth 2. The effect is not large
Donaldson (2015)	To examine what are the benefits of transportation improvements in trade and welfare	Railroads in colonial India: expansion between 1853-1930	Transportation infrastructure projects reduce trade costs, which, in turn, increase welfare	1. Ricardian trade model à la Eaton and Kortum (2002) 2. Lowest-cost route distance between two locations to model the cost of trading goods	OLS estimator for assessing the effects of railroads on trade costs, trade flows, real income levels and railroad impact	1. Railroads reduced the cost of trading and the inter-regional price gaps. They increased trade volumes, real agricultural incomes and welfare 2. Railroads allowed to benefit from gains of trade since their effect in reducing trade costs due to comparative advantage
Donaldson and Hornbeck (2016)	To explore the impact of the development of the railroad infrastructure on the American economy	Railroads in the U.S.: development between 1870-1890	Railroad network affects county's market access and enhanced market access is capitalised in agricultural land	Changes in county's market access to capture direct and indirect effects of railroads	OLS and two-stage least square estimators for assessing the effect of railroads (via market access) on the value of agricultural land	1. Railroads critical and irreplaceable to the agricultural sector 2. Railroads improve market access which, in turn increases land values
Fajgelbaum and Redding (2014)	To evaluate the impact of external integration on economic development and welfare	Big steamships and railroads in Argentina: improvement in late 19th century	Geographic heterogeneity within country affects external integration	1. General equilibrium model of the allocation of resources across regions and sectors 2. Shortest route as an instrument to identify connected locations	OLS and two-stage least square estimators for assessing the effect of railroads on agricultural productivity growth	Being close to world markets reduces relative price of traded goods, increases population, expenditure, demand of non-traded goods, relative price of non-traded goods, employment in non-traded sector
Jedwab et al. (2015)	To study the effect of historical shocks on the emergence and persistence of spatial equilibriums	Railroads in colonial Kenya: construction and expansion between 1896-1930	Historical shocks determine urban equilibriums which persists as consequence of path dependence	Distance from the main, branch, placebo railroads with a set of identification strategies to ensure causality of effects	OLS estimator for assessing the effects of railroads on past (population growth, European and urban inhabitants, agriculture, etc.) and present (population growth, nightlight luminosity, poverty) outcomes	1. Short-term: railroads affected European settlement pattern in Kenya. These settlements generated economic development which, in turn, determined railroads cities locations at independence 2. Long-term: railroads cities' economic development persisted as a mechanism to coordinate investments
Volpe Martincus et al. (2014)	To investigate the effects of road infrastructure on firms' exports and, then, on employment	Inca road network in Peru built by the Incas before 1530	Public investment in transportation infrastructure beneficial for firms' activity	Changes in current roads instrumented with the distance from the origin of exports to the nearest Inca road or the port travelled along the road	Instrumental variables approach on top of the difference-in-differences one for assessing the effect of new roads (via Inca roads) of firms' exports	Current road infrastructure has a positive effect on firms' exports. The increased exports lead firms to enlarge and to increase their labour forces
Berger and Enflo (2015)	To analyse both short- and long-run effects of transportation infrastructure on urban growth	Railroads in Sweden: first wave expansion between 1850-1870	Historical shocks help in explaining short- and long-run impacts on urban development due not randomly assignment across location of current infrastructure	1. Distinction between towns which gained access to railroads and non-connected towns. 2. Straight lines between endpoints as an instrument to identify connected towns	Difference-in-differences and instrumental variables approaches for assessing the effects of railroads on past and present urban population and other economic outcomes	1. Short-term: access to rail transportation, connecting towns, increased railroad cities' populations. However, this reflects relocation of industrial activity from non-connected towns 2. Long-term: historical railroad cities continued to grow also in following centuries

Source: Author's elaboration

Twelve years ago, Banerjee, Duflo and Qian started to investigate the long-term effects on economic growth of better access to transportation infrastructure. The interest was to empirically assess the causal impact on a twenty years period (1986-2006) of rapid growth in China. In order to avoid the potential endogenous location of roads, they rested on the idea of drawing a straight line connecting the historical Chinese cities as an instrument for the modern infrastructure (railroads and coastal routes) in China. They suggested that the straight lines linking the historical Chinese cities capture the 'old' historical Chinese railroads, constructed by the Chinese government in the late nineteenth-early twentieth century, which, in turn, capture the modern transportation infrastructure developed in China during the 1980s. Carrying out an innovative identification strategy and performing a detailed empirical analysis, the results emerging from their investigation confirmed how being close to transport networks leads, although moderately, to higher GDP per capita levels, increases the number of firms and has a positive effect on firms' profits. This work generated further investigations in following years by Atack et al. (2010) and Faber (2014) and assumed its final version in 2012. The straight-line instrument adopted by Banerjee and co-authors is particularly important in the discussion about the construction of Roman roads from an engineering point of view. Following Paragraph 2.6 and Chapter 3 will recall these thoughts.
The work of Atack et al. (2010) exploits the straight-line concept to evaluate the effect of access to railroads in the U.S. on urbanisation and population growth during the mid-nineteenth century. Drawing a straight line between two counties representing the starting point and the endpoint of a potential railroad, they use this instrument for predicting gaining rail access. However, differently from the contribution of Banerjee, Duflo and Qian, Atack et al. (2010) miss the link between contemporaneity and history. Similarly, in Faber (2014) this link is very frail. He exploits a straight-line instrument, based on Banerjee and co-authors' historical lines of transportation, in conjunction with other additional lines to capture the modern China's National Trunk Highway System (NTHS). Faber aims to evaluate whether the reduced trade costs resulting from the transport investments induce industrial concentration or diffusion of economic activity to periphery. He finds that the NTHS reduced the local GDP growth in non-connected peripheral counties due a reduction in industrial output growth.
Along the investigation on railroads, Donaldson in two subsequent works (the second one with Richard Hornbeck) analyses two great historical railroad projects: the railroad network designed and built by the British government in India between 1853 and 1930 and the railroad network expansion in the U.S. during the late nineteenth century.

In 2015, Donaldson, with its investigation on the impact of the Indian railroad network during the colonial India, presents three important contributions in the understanding of the effects of large transportation infrastructure projects and their expansion. First, he provides an estimation of the extent to which railroads improved India’s trading environment (reducing trade costs, lowering interregional price gaps, and increasing trade flows). Second, he assesses the welfare gains originating from the extension of the railroads. Third, he measures how many of these welfare gains are represented by gains from trade. Using a Ricardian model based on Eaton and Kortum (2002) in a general equilibrium framework and supposing that bilateral trade flows take the form of gravity equation, Donaldson is able to examine the effects of railroads on trade costs, on trade flows, on welfare, and what extent of welfare can be explained by the decrease of trade costs. He finds that railroads are the best instrument to reduce distance and trade costs. Trade costs increase with effective distance and the railroad freight rates are lower than those of roads, rivers or coastal transport. When considering trade flows, railroads significantly enlarged trade in India, producing also welfare gains due the increased gains from trade. Moreover, in terms of welfare gains, the access to the railroad network raised the real agriculture income by over 16 percent. The entire welfare gains come from trade costs and trade gains.

The work of Donaldson and Hornbeck of 2016 examines the impact of the historical late nineteenth century railroad network on the American economy, focusing on the agriculture sector. The expansion of the railroad network in the U.S. during the late 1800s represented an important change for the development of the country. Railroads became the principal way of transportation and territories close to railroads experienced an increase in terms of income and wealth. In order to assess the effect of railroads on the U.S.' economy, the authors consider the impact on the counties' market access. The focus on agriculture sector and on counties' market access was legitimated, as the authors pointed out, by its empirical advantage. The county's market access increases when trading with another county becomes cheaper and when the other county has a larger population or faces higher trade costs with other counties. Counties' market access are influenced by alterations elsewhere in the railroad network. Moreover, as the railroad network extended, changes in market access were capitalised into agricultural land values. Donaldson and Hornbeck (2016) find that railroads had a positive effect on the agriculture sector. Removing all railroads access in 1890 would reduce the Gross National Product (GNP) of 6.3 percent, which roughly doubles the estimation by Fogel (1964) of 2.7 percent, who argued that the railroad network in the United States was less effective than canal networks during the same period.

Two further works that look at historical railroads are the contribution of Jedwab et al. (2015) and the paper of Berger and Enflo of 2015.

Jedwab and co-authors put their work in that vast literature aimed at understanding what determines the distribution of economic activity across space and how historical shocks shape the pattern development. In Subparagraph 2.2.3 some of these contributions (Davis and Weinstein, 2002; Miguel and Roland, 2011; Redding et al., 2011) have been examined; however the empirical evidence is various. The work of Jedwab et al. (2015) analyses the colonial railroads in Kenya as a natural experiment to study whether history generates spatial equilibrium that is feasible and persists in time. The colonial period in Kenya started in 1895 and ended in 1963, when it became independent from the British government. During that period, in 1896, European colonisers began the construction of the railroad in Kenya, that reached its maximum extension in 1930. Jedwab, Kerby and Moradi focus on this year to construct their measure of railroads as different distance dummy variables (i.e. one dummy for each range of distance in kilometres). Exploiting both rail and ‘placebo’ (railroads that were planned but not constructed) lines, they perform a double empirical analysis. On the one hand, they assess the effects of railroads in the short-term, using as dependent variable outcome variables related to the colonial period: population growth, European inhabitants and agriculture, and other historical factors. On the other hand, Jedwab et al. look at the long-term effects, exploiting long-run and modern economic measures (population growth, nightlight luminosity, poverty, and other contemporary factors). The authors find that railroads had a main role in determining the outline of European settlements which, in turn, due economic development they generated, influenced the location of Kenyan cities at independence. These cities are the more developed still today, confirming the persistent effect of railroads. Regarding the matter of industrial distribution across space, Jedwab, Kerby and Moradi confirm the existence of path dependence and multiple steady states for economic activity locations, as suggested by Redding et al (2011).

The contribution of Berger and Enflo (2015) is in line with conclusions emerging from Jedwab et al. (2015). Looking at the railroad infrastructure in Sweden, built during the middle of the nineteenth century, Berger and Enflo aim to explore the effects of transportation infrastructures on urban growth and location of economic activity across towns, relying on the fact that historical shocks, given by improvements in the access to transport infrastructure, help to overcome the non-randomness of current transportation networks. They distinguish towns that had access to railroads during the first wave (1850-1870) of railroad construction (treatment group) from non-connected towns (control group). Moreover, they connect endpoints by straight lines and measure a buffer zone, creating an instrument to identify connected towns. Exploiting both a difference-in-differences and an instrumental variables approach, they find short- and long-term effects. Railroads increased urban population in towns that gained access to the transportation infrastructure. This growth is mainly due to a redistribution of industrial activities from non-connected cities to connected ones. The increasing in population lasted in the long-run: cities which grew during the first wave of construction expanded in later centuries too. Finally, Berger and Enflo contribute to the literature on the location of economic activity and existence of a unique equilibrium or multiple steady states, arguing that path dependence in the distribution of industrial activities occurs.

Fajgelbaum and Redding in 2014 look at Argentina and at the reduction of international transport costs generated in late-nineteenth century by the introduction of big steamships. The integration into world markets of Argentina is used as natural experiment to examine how geographic heterogeneity within country affects external integration which, in turn, determines economic growth and welfare. The improvement in sailing reduced the external shipping cost, making it profitable to export from Argentina to Europe a large amount of agricultural products and meat, and stimulating the expansion of the railroad infrastructure. Developing a general equilibrium model and exploiting the same model for the quantitative analysis, for which also an instrumental variables approach for the rail infrastructure has been developed, Fajgelbaum and Redding highlight how being close to world markets increases population density, employment, relative prices of non-trade goods and of land.

Last but not least, the original contribution of Volpe Martincus, Caraballo and Cusolito concludes this review. Volpe Martincus et al. (2014) analyse the case of Peru, in order to assess which has been the effect of the new road infrastructure on firms' exports and employment and to investigate whether public policies supporting transportation infrastructure projects positively affect firms' global trade. Peru has experienced in the first decade of two thousands an important expansion of its road network and some figures claim that firms, in those municipalities whose roads connections to the main port improved, increased both exports and employment between 2003 and 2010.[3] Although this statistical evidence, it is not totally informative due the problem of reverse causality that typically affects the relationship between infrastructure and exports or infrastructure and

[3] More than 5,000 kilometres of new roads were constructed in Peru between 2003 and 2010 (Volpe Martincus et al., 2014).

employments.[4] In order to not incur in this endogeneity concerns, Volpe Martincus at al. use the Inca road network (built by the Inca empire before 1530) as an instrument for the current road infrastructure in Peru. The Inca road network represents an exogenous source of variation in transport infrastructure, since the reasons behind its construction are completely unrelated with current firms' foreign trade and employment. Moreover, the authors confirm that the only channel through which the Inca roads affect today's exports is represented by the correlation with the spatial allocation of new roads. Instrumenting the change in road infrastructure between 2003 and 2010 with two instruments (the distance from the geographical origin of the exports to the nearest road that was part of the Inca network, and the distance between this origin and the current port that could have been travelled along roads in this network), they find that domestic road infrastructure increased firms' exports, and this expansion in global trade induced firms to hire more employees. Moreover, improvements in road networks lead advantage to firms' exporting activity and job growth, thanks to the ease of more and larger shipments that the new roads allowed.

2.5 The persistent effect of Roman roads

The economic history has widely devoted interest to the magnificent Roman past of most of European countries, emphasising how the Roman empire has represented for more than seven centuries an interesting and impressive domination from an historical point of view as well as from an economic perspective. During recent years, rather than investigating the past, the economic literature started wondering whether the Roman legacy persisted until today. This interest has become significant and the greatness of the empire, and all those elements characterising the Roman domination, have started to be included in appealing contributions focused on the understanding of the long-term effects of the Roman world. In some of these works the importance of the Roman road infrastructure is mentioned, but it has been only during last years that the literature has begun to investigate on the persistent effect of Roman roads on modern economy and, at the time of writing this thesis,[5] there are only other two works[6] completely devoted to analysing whether there has been and what has been the impact of the Roman road network on present economic outcomes.

In the work of Buringh et al. (2012), the Roman road issue appears marginally: the road stations are used to identify a type of settlement in Roman times. Buringh and co-authors aim

[4] Paragraph 3.2 in Chapter 3 presents a review of the endogeneity problem which typically affects the relationship between infrastructure and economic activity.

[5] October 2016.

[6] Wahl (2015) and Dalgaard et al. (2015).

to study the development and the spatial distribution of a market economy in Roman times using the size and place of Roman settlements as a proxy for the urbanisation and the distributions of low denomination coins as a proxy for the market economy. They find that the Roman military presence in a given area and the big demand of goods and services is the explanation behind the constitution of a market economy and the differences in the economic development in north-western Europe around the year 150 A.D. These differences are confirmed also by the recovery of coins. The presence of a market economy can be retrieved in the small denomination coins circulating during the Roman period; the absence of these coins during pre-Roman times and the higher fraction of precious metals in the coins found after the collapse of the Roman empire confirm that the market economy has developed with the military payments, which not only played an important role in creating a cash based market economy, but contributed to urbanising those territories where the military demand of goods and services was high. The perspective of the paper is mainly devoted to understand the past, rather than investigating the effects on the present, and the use of the Roman road subject is only instrumental. On the other hand, the work informs about how originally military settlements had later on developed into a mixed civil-military settlement by the arrival of civilians, highlighting a legacy of the Roman domination in influencing subsequent developments and urbanisation.

Bosker et al. (2013) use Roman roads as a potential for cities to trade. The contribution is aimed to appreciate the different development of Europe and of the Islamic world in a long-run period, from 800 to 1800. The authors find that the reasons behind Europe's growth and Islamic world's stagnation must be retrieved in specific characteristics of the two worlds, since the interaction between them was limited by their divergent religious orientation. In the work, Roman roads are used to identify those settlements that could benefit from a close land-based transportation. According to the authors, the Roman roads have the advantage of being uniform in both Europe and Islamic world, since they were constructed with similar standards and methods in the entire Roman empire, and they are not affected by endogeneity. Once more, the Roman road reasoning is used as a tool to identify those settlements that could benefit from the proximity to the road, but differently from the previous work, there is no interest for the Roman world. On the other hand, the authors remark how the fall of Roman empire caused a period of stagnation in Europe and suggest that the role of the Roman road network in determining development and urbanisation in the Islamic world was limited and weak.[7]

While Roman roads played a marginal role in influencing settlement in North Africa and in the Middle East, this is less the case when considering Europe. A previous work of Bosker and Buringh (2012) is aimed at understanding what determined urbanisation in Europe over the period 800-1800. The authors find that geographical features (like being close to water- or land-based transportations, the existence of agricultural possibilities, accessibility) have been fundamental in determining the city location during the starting phases of the constitution of the European urban centre. It was only in a second stage, during the seventeenth century onwards, and with the decrease of trade costs, that the geographical characteristics of the surrounding urban system (like being close to already existing cities) influenced the city location. In this contribution, rather than appreciating the comparison between two different worlds (the European and the Islamic), the attention is completely devoted to the old continent, and the Roman road issue comes out in a more substantial manner. Bosker and co-author, as in the latter work (with J.L. van Zanden), use the Roman roads to identify those locations that take advantage from being located close to a road or, even, where two or more Roman roads crossed. The findings coming from the work of Bosker and Buringh (2012) are interesting. Their results suggest that both location on a river and by the sea significantly and positively affect the possibility of a location becoming a city. Whereas, "*Good location for land based transportation instead has a surprising negative effect. Locations on the former roman road network have lower urban chances (even when located on a hub of two roman roads)*" (Bosker and Buringh, 2012, p.17). These results are not robust when distinguishing the one thousand years period between pre-1600 and post-1600. In the pre-1600 period, being located at a hub of Roman roads is beneficial for a settlement; in the post-1600, this advantage becomes insignificant or even negative (disadvantage).

The work of Micheals and Rauch (2014) represents one of the first attempts in studying the impact of the Roman empire after two millennia. In particular, Micheals and Rauch analyse whether history is able to affect and to entrap the location of a town for centuries or whether geographical features are the only ones responsible for it. In order to test this research question, they trace the locations of towns in France and Britain back to the Roman Empire. France and Britain represent a perfect case study: they were both touched by the Roman

[7] As highlighted by Wahl (2015) the modest or absent effect of Roman roads in fostering development in North Africa and the Middle East may be explained by the fact that the "*Roman road network was not very dense in Northern Africa and in general, the course and existence of the roads is uncertain*" (Wahl, 2015, p. 3).

domination and, in later centuries, French and English towns were characterised by similar institutions, organisation and size. But, after the collapse of the Western Roman Empire, France and Britain went through different destinies. While in Britain the urban network stopped, in France this was not the case. The fall of the Western Roman Empire represents the shock they study to investigate the spatial consequences of the relocation of towns. The authors argue that, from the Early Middle Ages until the beginning of the Industrial Revolution, French towns, compared to the British ones, were settled three times closer to old sites of Roman towns. This medium-term persistence is confirmed in nowadays France and Britain's urban locations, suggesting how town settlements in Britain relocated more than in France and how urban networks are affected by path dependence not only in the medium-run, but in also in the long-term. In the work of Micheals and Rauch, the Roman road issue is used to explain whether the difference between the urbanisation in France and Britain can be ascribed to the road network. Roads were constructed to let the move of the Roman army in a quick and easy way, and during the hegemony of Roman empire some towns grew alongside the roads. But during the Middle Ages, with the deterioration of road quality and with the increasing improvement of water transportation, coastal access became more important and towns were more prone to having a port.

In all these studies, the heritage of the Roman empire comes out in the development and urbanisation in future periods: most of old Roman towns became important cities in subsequent centuries[8] and the Roman legacy appears more acting in the medium-run (during the Middle Ages) rather than in the long term. Moreover, from these contributions, it emerges the important role of roads in facilitating settlements and in creating urbanised centres[9] in early stages of urbanisation. Although the key function of roads comes into view, none of these works is entirely focused on understanding which has been the legacy of the old Roman road network. The works of Wahl (2015) and of Dalgaard et al. (2015)[10] are the first ones, together with Chapter 3 of this thesis, that study whether the Roman road system is able to affect the modern economy by its persistent effects.

Wahl (2015) argues that the distinction between a West developed Germany, crossed by an integrated Roman road system, and a less advanced (compared to the West) East Germany originates from the division in ancient times between a Roman and a non-Roman part. The Roman Western and Eastern parts of Germany do not correspond with the division between

---

[8] See Pirenne (1969) and McCormick (2001) for a more detailed analysis.

[9] Lopez (1956) discusses the importance of the Roman road system in Europe during the Middle Ages.

[10] At the time of writing the thesis (October 2016), there isn't any working paper accessible. The review is based on the slide presentation available at: http://economics.gu.se/digitalAssets/1531/1531740_roman4.pdf.

West and East Germany after the World War II, but the separation between the two areas corresponds with the track of the German Limes wall, which crosses the entire South part of Germany and the most Central-Western Germany. In fact, differently from other countries (like France, Italy and Spain), Germany has not been completely touched by the Roman domination and the Limes wall represents the geographical discontinuity between the Roman German part and the non-Roman one. Wahl takes advantage from this fact and bases the empirical identification of an effect of the Roman road network on a spatial regression discontinuity approach using the German Limes border as a cut-off to separate treated (crossed by Roman roads) and non-treated (not crossed by the Roman infrastructure) areas. Using the nightlight luminosity, Wahl shows that the Roman Limes border wall represents the boundary in the economic development of actual Germany. Moreover, Wahl shows how the Roman road infrastructure positively influenced the constitution, the growth and the agglomeration of cities resulting in current higher urbanisation and economic progress.

Likewise, employing the current nightlights intensity and the population density in Europe, Dalgaard et al. (2015) find a long-run causal effect of Roman roads on economic development. The intensity of the Roman road network is able to predict the historical urban population levels and current levels of economic activity. According to Dalgaard and co-authors, the Roman road construction process represents an ideal natural experiment for three main reasons. First, Roman roads were not easily useable by travellers: they were straight in order to economise on costs. Second, Roman roads were built primarily for military purposes. Third, Roman roads were constructed on populated but undeveloped territories.

2.6 The Roman road network: a brief overview

The Roman empire has represented, in Italian and non-Italian history, one of the (say the) greatest empires of all time in terms of territory possessed and duration of political power. As highlighted by Laurence (1999), historians have recognised that the Roman state was involved in the development of an extensive transport network of roads from the fourth century BC, but have not managed to understand the impact of road building.

The tactical purpose, the logistics for the war campaigns and the supply of the army across roads represent the spirit of the road system and of the whole Roman empire. Although the military aim is at the core of the road network, when talking about Roman roads there are other four main aspects that should be considered and that are strictly related with the constitution of an intricate road system: i) the development of a great empire, ii) the advanced engineering abilities of the Romans, iii) the openness and mobility culture, iv) the other ways of transport and trade. In this paragraph, all five themes (military, empire, engineering, openness, transport) will be briefly covered, focusing on how the army purposes led the engineering knowledge of Romans to construct roads that lasted until today, that served the foundation of a complex and huge empire continually expanding and that enabled openness and connectedness.

Starting from the military theme, Roth (1999) underlines that the Romans built their roads primarily for military reasons; the commercial travel was only an indirect advantage of the road network. The majority of the weight of the Roman army supplies was represented by three elements: food, fodder and firewood. Hence, all military decisions were determined by the need to assure the provision of supplies to the army. Accordingly, Thompson (1997) explains that the construction of a road network originated from the need to ensure that a large number of horses, cattle's, carts and infantry could circulate: primordial non-Roman routes were problematic during wet and rainy days, since the deep mud impeded or delayed the movement of goods and services. Therefore, the construction of a network of paved roads empowered not only the transport of goods and services, enabling the movement of larger quantities and people and making transfers easier, but armies were able to travel twenty-five miles a day, even in bad weather conditions. During the Republican and Imperial periods, the Roman empire conquered territories in the Mediterranean Sea (like Sicily, Sardinia, Corsica, the northern coasts of Africa), in the Atlantic Ocean and in the Black Sea. Roth (1999) describes that the Romans were aware that moving supplies to the army by ship was cheaper and faster than by land, but transport by sea was dangerous and expensive. Seafaring in the Mediterranean was limited between March and November, although it was really safe only during the summer months (from June to September). On the contrary, land transport had no limitations and was practicable all year round. The Romans understood that the need of a logistical infrastructure was fundamental for the movement of armies and supplies and consequently for the enlargement of the empire. Hence, in order to facilitate the move of the army from place to place, they constructed roads intended mainly for wagon travel used for supplies; soldiers and pack-animals could travel as well. Since the military purpose was the priority for the Romans, expansion, maintenance and repairing of the road network were continuously performed for and associated with military campaigns and strategic and tactical purposes.

Historians have argued how the design of an intricate road system and the development of a great empire were strongly correlated. On this point, Thompson (1997) argues that the vast and comprehensive Roman road system changed the entire empire. Accordingly, Gleason (2013) stresses that the enlargement of the empire was possible thanks to a developed road system. In fact, the Roman army was too small to conquer the enemies of Rome, but the constitution of a vast empire was possible investing in the construction of a complex road network rather than enlarging the infantry forces. Knapton (1996) underlines that the new conquered regions, on the one hand, enlarged the Roman empire contributing to its power, authority and wealth; on the other hand, the payments of conquered territories were mainly used for the public infrastructures, like roads and aqueducts. This led to the development of the engineering capability of the Romans. The peak of the Roman empire corresponds with full extent of the road network (117 A.D.-death of Trajan) as further proof of the fact that roads construction and the constitution of a vast empire were highly related.

The engineering behind the construction of roads subtend incredible and high-level skills. Romans were mainly focused on getting the road straight, since it was easier for the network structure and shape. To achieve this straight configuration, they defined points that could be quickly connected by a straight line (Davies, 1998). Accordingly, Gleason (2013) explains that to mark the road's path with either stakes or furrows, creating roads as straight as possible, was the first purpose of Romans. Legionaries and slaves belonging to the army were involved in the roads construction process. This process included first of all the digging of a 1.5 meters deep trench for the width of the road. In order to guarantee the stability and durability of the substrate, the trench was filled and packed with several texture and types of material from the land around it. Then they applied a layer of gravel or pavestones, ensuring that the road had a camber, or rise in the centre, to prevent erosion and make the surface all-weather capable.

Behind the development of an intricate and technological road system there is a culture of openness and mobility. Geographical distance between places creates distance between people. Knapton (1996) describes how the Roman road network represented a system which connected different peoples and cultures from Newcastle to North Africa, from Portugal to Arabia. The Italian territory was itself a mixture of peoples: Greek colonies and the Samnites in the south, the Latins and the Sabines in the centre, the Etruscans in the north of Rome, the Celts in northern Italy and other peoples in the rest of the peninsula. During what has been called 'the golden age of Roman road building' (second century B.C.), the Roman empire became interlinked with a network of roads which led to greater mobility and the Romans used to live overseas and to become wealthy. On this point, Laurence (1999) highlights that the understanding of the nature of Roman space-time is fundamental to appreciate the cultural change associated with road building and the improvement in terms of road technology. Roman roads changed the speed of communication and created connections throughout the year: the space-time concept integrated the elements of physical distance and time taken to complete a journey over that distance. The road system created an interconnection between places that allowed for a mobile elite and citizen body and also the mobility of surplus products and profits. The developments in the road system technology and the increased efficiency of transports, together with a state of mind of space-time that emphasised the transport of people and goods over a distance, were features of a culture that had an emphasis on mobility. The issue of mobility was embedded particularly in a system of elite land holding, that depended on mobility of the landowner for its economic survival, and an elite culture, that laid claim to active involvement in the management of their estates (Laurence, 1999). Accordingly, Sommer (2015) suggests the big role played by the Roman road system in making goods, people and ideas mobile as never before.

Road transport can be seen as a complementary system of river and maritime transport. It has been argued that land transport was an inferior, expensive alternative to maritime transport. Goods were constantly transported throughout the Roman Empire and it has been largely discussed that, despite the risks, dangers and problems, the most efficient way to transport goods was by sea. According to its size, it could carry cargo weighing between 70 and 350 tons (Snedden 1998). Ships were preferred to roads since they could transport large amount of goods and people in a shorter time. Six hundred passengers or six thousand amphorae of wine, oil and other products were highly traded using sea transports. The Romans put much effort in improving the effectiveness of shipping, developing harbours and lighthouses. The journey from Egypt to Rome took only two to three weeks by ship. When transporting commercial goods, river transport was also used and preferred to roads, and the same principle applied for the movement of military supplies. The access to the inland regions of the empire was allowed by the large navigable rivers: the Rhône into Gaul, the Rhine into Germany, the Danube into Pannonia, Dacia and Noricum, the Tigris and the Euphrates into Mesopotamia and the Nile into Ethiopia (Roth, 1999). Despite the apparently overwhelming economic advantage of trade by water, Pawson (1977) pointed to the key advantages of land transport, arguing that the land transport system could be classified in two parts: a complementary system, which was interdependent with water transport and performed a feeder and distribution role for it, and a competitive, independent system which did not rely on water transport linkages.

2.7 The Roman road network: the starting data set

Geo-coded data on the Roman road network refers to the McCormick et al. (2013) shape file, a linear layer obtained digitizing the information in the "Barrington Atlas of the Greek and Roman World" (2000). Data is made freely available on the internet by the Digital Atlas of Roman and Medieval Civilizations (DARMC), which allows spatial and mapping analysis for the Roman and medieval worlds using the Geographic Information Systems (GIS) approach. The level of geographical accuracy and detail and the georeferenced data allow spatial analyses and careful investigations. Moreover, the strict connection with the work of Talbert[11] confirms the reliability and the correctness of information included.

A general picture about the extension and the branch of the Roman road network and a representation of the structure of the data through some descriptive statistics and some maps is fundamental to better appreciate the magnitude and the importance of the Roman road infrastructure. On this aim, Figure 2.1 shows the data set as provided by the DARMC.

Figure 2.1 Roman Roads at peak of empire (117 CE) - McCormick et al. (2013) data set

Source: Author's elaboration from McCormick, M. et al. 2013. "Roman Road Network (version 2008)", DARMC Scholarly Data Series 2013-5

[11] Barrington Atlas of the Greek and Roman World (2000).

It consists in a shape file including 7154 segments of ancient Roman roads existing at the peak of the empire, corresponding with the death of Trajan (117 CE). The network covers 36 countries[12] over Europe, Africa and Asia. Roads are classified if major or minor and if certain and uncertain[13] roads. The peculiarity of the shape file is represented by the way according to which roads are recorded: it does not classify single and complete roads, but it provides a list of segments which compose roads.[14] In other words, roads are split into portions, and for each portion the information about its length in metres, its size and its certainty is made known. Some historians disagree with this clear categorization of roads, arguing that some of those roads classified as uncertain in the data set of McCormick et al. (2013) are instead certain, since the recovery of Roman milestones in those territories, marked as areas crossed by Roman roads, is an unambiguous sign of the existence of the road. The empirical application of Chapter 3 took advantage from this puzzle, performing the empirical analysis using only certain roads, only major roads, both certain and major roads together and all roads included in the data set.

Size and certainty of the road are the two main features evaluated in the shape file. Figure 2.2 and Figure 2.3 provide a representation of these two characteristics classifying roads between major and minor and between certain and uncertain, respectively. From the observation of Figure 2.3, it emerges that there is a reasonable balance between certain (orange lines) and uncertain roads (red lines).[15] This is less the case when considering the size of the road: Figure 2.2 shows a higher presence of blue lines (minor roads) than red ones (major roads). Moreover, when drawing only major roads, it is immediately possible to distinguish present nations' borders and to recognise the outline of the Mediterranean Sea, since a high number of roads have been constructed along the coasts.

The total Roman road network is almost 200 thousands kilometres long.[16] The information on the length of each segment of the Roman road network is the third trait considered in the data set. It originates from the typology of the data set (geo-coded data). On the one hand, it allows to appreciate the extent and the coverage of the whole Roman infrastructure. On the other hand, it does not allow to appreciate the length of the infrastructure by geographical

[12] Albania, Algeria, Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Egypt, France, Germany, Greece, Hungary, Israel, Italy, Jordan, Lebanon, Libya, Liechtenstein, Luxembourg, Macedonia, Montenegro, Morocco, Netherlands, Palestine, Portugal, Romania, Saudi Arabia, Serbia, Slovenia, Spain, Switzerland, Syria, the U.K., Tunisia, Turkey.

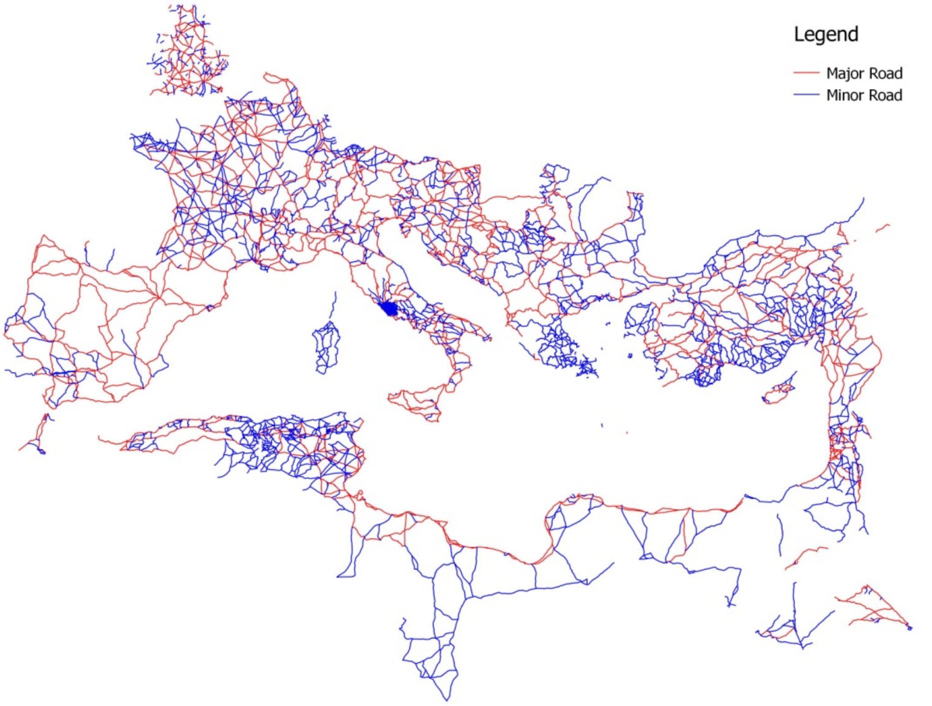
[13] The certainty refers to the path followed by the road. The road is certain in the existence and in the Roman origin, it is uncertain in the route.

[14] Each segment has a unique id number.

[15] For the majority of Roman roads in Greece and Turkey there is no information available about the certainty.

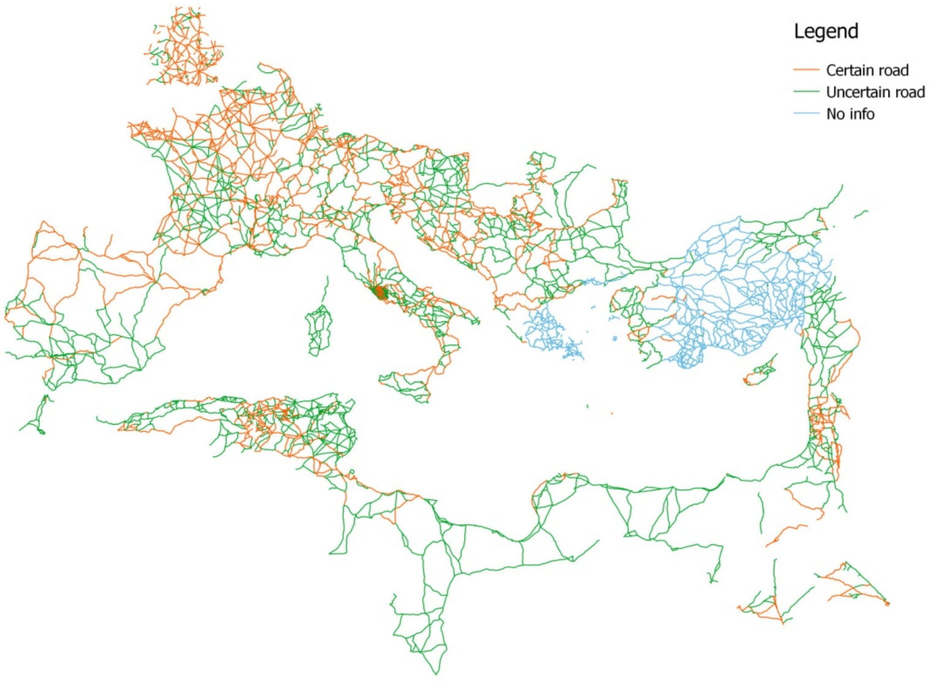
[16] Exactly 192,861 kilometres.

Figure 2.2 Roman roads by size: major and minor roads



Source: Author's elaboration from McCormick, M. et al. 2013. "Roman Road Network (version 2008)", DARMC Scholarly Data Series 2013-5

Figure 2.3 Roman roads by certainty: certain and uncertain roads



Source: Author's elaboration from McCormick, M. et al. 2013. "Roman Road Network (version 2008)", DARMC Scholarly Data Series 2013-5

territories. As remarked by the DARMC, the data set of the Roman road network is thought to be used in combination with own historical data. In fact, the potential of the DARMC's shape file lies more on what can be done with it rather than on the mere representation of the Roman road system. In other words, the use of geographic information system approaches and the implementation of technical and geometric tools to the starting shape file allows researchers to extract new information and to generate new data ad hoc devised for specific investigations. This will be at the core of the following Paragraph 2.8, which describes what has been performed and what has been produced starting from the Roman road layer provided by the DARMC.

2.8 The Roman road network: the new-created data sets

The precious information contained in the DARMC's shape file, the potential elaborations it allowed using GIS instruments and the promising research question of following Chapter 3[17] have inspired the creation of two new data sets: a shape file of the Roman road infrastructure for the sole Italian territory and a measure of Roman roads in kilometres by Italian province. Moreover, since the computation of the Roman road measure in kilometres has originated from layers, 108 layers for each Italian province that includes Roman roads have also been created. On the one hand, the new constructed data sets should be considered as a contribution to that strand of research aimed at studying and collecting all those elements and features of the Roman world in order to create new and useable data sources. On the other hand, the method exploited to construct them provides a scheme and a technique that can be easily converted for other investigations and simply followed and reproduced for extended or reduced territorial analyses.

Figure 2.4 shows the newly created layer of the Roman road system for the sole Italian territory, obtained applying a specific geoprocessing tool to the starting shape file provided by the DARMC. Italy includes 25 percent of the entire network, where, unsurprisingly, Roman roads are more concentrated. Moreover, the provincial partition displayed in Figure 2.4 allows to appreciate the ramification at the NUTS3 level. Since the Italian shape file is a subset of the DARMC's layer, Table 2.3 allows to compare the two shape files in terms of number and types of segments. The complete starting data set is composed by 7154 parts of Roman roads; 1817 when considering just Italy. 43 percent of the complete data set are major roads, in Italy they represent 35 percent of the total Italian segments. Almost half of the segments are

[17] Chapter 3 presents an empirical investigation on the persistent effect of the Roman road network on trade costs of Italian provinces.

Figure 2.4 New created Roman road layer for Italy

Source: Author's elaboration from McCormick, M. et al. 2013. "Roman Road Network (version 2008)", DARMC Scholarly Data Series 2013-5 and from Istat data (2011)

evaluated as certain in both complete and Italian layers. When considering both features (size and certainty), the majority of segments are minor and uncertain. Differently from Table 2.3 that describes the segments composing the whole and the reduced (Italian) data set, Table 2.4 takes the point of view of Italian provinces. 108 out of 110 provinces have Roman roads,[18] 94 have major roads, 85 have certain roads, 77 have both certain and major Roman roads.

The creation of the new Italian layer has represented the first stage for the construction of the new measure of Roman roads in kilometres by province. The computation has consisted in steps[19] and has been performed using GIS and two digital maps: the linear shape file of Roman roads for the Italian territory and a polygonal shape file of 110 Italian provinces provided by Istat.[20]

[18] The two Italian provinces where the Roman road network is absent are the province of Pordenone (in the north-east) and the province of Verbano-Cusio-Ossola (in north-west).

[19] Appendix 2.1 provides a detailed explanation on how the measure has been obtained.

[20] Istat (Istituto Nazionale di Statistica) is the Italian National Institute of Statistics that provides the official statistics for Italy.

Table 2.3 The "Roman Road Network (version 2008)" shape file description: complete and Italy layer

	Number of segments	
	Complete starting layer	New Italy layer
Size of the road		
Major roads	3144	640
Minor roads	4010	1177
Total	7154	1817
Certainty of the road		
Certain roads	3117	879
Uncertain roads	3131	938
Not defined	906	-
Total	7154	1817
Size and certainty of the road		
Major and certain roads	1911	466
Major and uncertain roads	1006	174
Major and undefined certainty roads	227	-
Minor and certain roads	1206	413
Minor and uncertain roads	2125	764
Minor and undefined certainty roads	679	-
Total	7154	1817

Source: Author's elaboration from McCormick, M. et al. 2013. "Roman Road Network (version 2008)", DARMC Scholarly Data Series 2013-5

Table 2.4 Features of Roman road segments by Italian province

	Number of provinces
Size of the road	
Only major roads	20
Only minor roads	14
Both major and minor roads	74
Total	108
Certainty of the road	
Only certain roads	9
Only uncertain roads	23
Both certain and uncertain roads	76
Total	108
Size and certainty of the road	
Major and certain roads	77
Major and uncertain roads	78
Minor and certain roads	41
Minor and uncertain roads	79

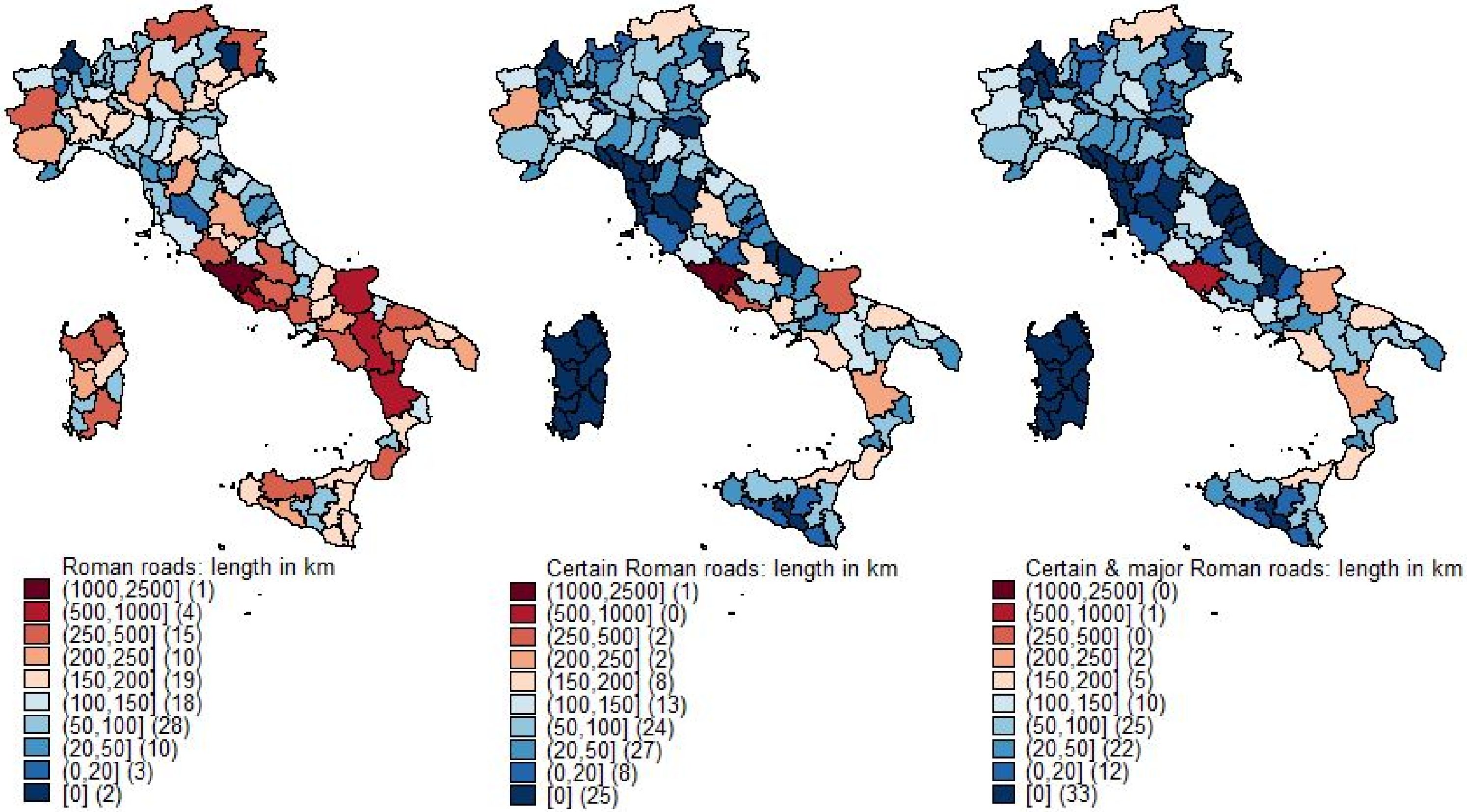
Source: Author's elaboration from McCormick, M. et al. 2013. "Roman Road Network (version 2008)", DARMC Scholarly Data Series 2013-5

Table 2.5 Kilometres of Roman roads by Italian province

Italian Provinces	All Roman roads: length in km	All Roman roads: rank	Certain Roman roads: length km	Certain Roman roads: rank	Italian Provinces	All Roman roads: length in km	All Roman roads: rank	Certain Roman roads: length km	Certain Roman roads: rank
Roma	2498	1	1125	1	Ravenna	124	56	66	41
Latina	671	2	264	3	Crotone	117	57	49	52
Foggia	570	3	397	2	Mantova	116	58	51	50
Potenza	560	4	119	19	Aosta	113	59	113	23
Cosenza	531	5	249	4	Livorno	112	60	0	86
Salerno	427	6	169	9	Grosseto	110	61	4	84
Frosinone	386	7	91	31	Chieti	109	62	0	86
Olbia-Tempio	379	8	0	86	Modena	107	63	28	70
Torino	366	9	211	5	Genova	105	64	78	35
Caserta	356	10	166	12	Pesaro-Urbino	103	65	103	26
L'Aquila	342	11	187	6	Parma	101	66	49	51
Bari	280	12	179	8	Savona	100	67	53	47
Viterbo	272	13	114	22	Lodi	100	68	8	81
Udine	271	14	126	15	Como	99	69	69	40
Bolzano	266	15	168	10	Caltanissetta	98	70	0	86
Palermo	265	16	96	28	Enna	93	71	9	80
Matera	261	17	53	46	Vercelli	91	72	41	56
Sassari	261	18	0	86	Isernia	90	73	31	66
Cagliari	255	19	0	86	Lecco	89	74	32	64
Reggio Calabria	251	20	185	7	Carbonia-Iglesias	86	75	0	86
Oristano	250	21	0	86	Ogliastra	86	76	0	86
Avellino	245	22	44	54	Ferrara	78	77	0	86
Perugia	238	23	155	13	Ascoli-Piceno	77	78	20	78
Agrigento	235	24	6	82	Rovigo	76	79	22	76
Firenze	229	25	39	59	Bergamo	76	80	40	58
Brescia	228	26	96	29	Arezzo	71	81	0	86
Lecce	210	27	32	65	Vibo Valentia	70	82	24	75
Verona	206	28	120	18	Forlì-Cesena	69	83	52	48
Taranto	202	29	97	27	Medio Campidano	69	84	0	86
Cuneo	200	30	72	39	Pisa	65	85	0	86
Campobasso	200	31	36	61	La Spezia	65	86	0	86
Treviso	198	32	117	21	Reggio Emilia	62	87	62	42
Pavia	192	33	118	20	Asti	59	88	52	49
Trapani	188	34	41	57	Ancona	59	89	59	45
Bologna	184	35	110	24	Belluno	57	90	48	53
Nuoro	182	36	0	86	Sondrio	56	91	11	79
Brindisi	176	37	123	17	Pescara	54	92	0	86
Siracusa	176	38	85	33	Massa Carrara	54	93	0	86
Venezia	176	39	29	69	Novara	52	94	25	74
Catanzaro	175	40	92	30	Vicenza	52	95	36	62
Catania	175	41	83	34	Rimini	49	96	29	68
Padova	173	42	37	60	Lucca	48	97	0	86
Messina	167	43	167	11	Trieste	44	98	27	72
Piacenza	166	44	125	16	Gorizia	41	99	41	55
Alessandria	164	45	126	14	Imperia	34	100	34	63
Ragusa	162	46	4	83	Pistoia	32	101	0	86
Cremona	161	47	74	37	Varese	27	102	27	71
Terni	153	48	78	36	Fermo	27	103	27	73
Benevento	151	49	61	43	Monza	23	104	0	86
Milano	148	50	74	38	Macerata	20	105	20	77
Trento	138	51	89	32	Siena	18	106	0	86
Barletta-Andria-Trani	137	52	108	25	Prato	10	107	0	86
Teramo	134	53	29	67	Biella	7	108	0	86
Napoli	129	54	60	44	Pordenone	0	109	0	86
Rieti	127	55	2	85	Verbano-Cusio-Ossola	0	109	0	86

Source: Author's elaboration from McCormick, M. et al. 2013. "Roman Road Network (version 2008)", DARMC Scholarly Data Series 2013-5

Figure 2.5 Roman roads in length by Italian province



Source: Author's elaboration from McCormick, M. et al. 2013. "Roman Road Network (version 2008)", DARMC Scholarly Data Series 2013-5 and from Istat (2011)

According with the information about size and certainty provided, five different measures have been computed:

- kilometres of all Roman roads;
- kilometres of just certain Roman roads;
- kilometres of just major Roman roads;
- kilometres of both major and certain Roman roads;
- binary variable: 1 if the province has Roman roads.

In order to have a road density measure, the first four measures have been weighted by the area of the province in square kilometres (Source: Istat) in order to have a general picture about the density of the road system.

Table 2.5 ranks Italian provinces according to the length in kilometres, considering all and certain Roman roads. Unsurprisingly, the first position is occupied by the province of Rome which owns, in absolute terms, the largest extent of roads, both considering all Roman roads and just certain roads. Provinces of Latina (close to Rome) and Foggia (in Apulia, south-east) are in second and third position. The picture which emerges from the table is that the Romans built roads mainly in the central and south-eastern part of the peninsula: the first fifteen places in both classifications comprise mostly southern provinces; only four are in the north: Torino, Udine, Bolzano and Alessandria. The 'numerical' information provided by Table 2.5 is confirmed by Figure 2.5. Figure 2.5 shows the distribution of Roman roads in the Italian peninsula. When considering all the roads (certain and uncertain, major and minor), it seems that the Romans devoted more efforts to building roads in the south rather than in the north. A possible explanation of this fact can be found in the way the Roman empire enlarged: firstly towards the southern regions, then to the north. When looking just to the certain or to certain and major roads together, this weak spatial distribution disappears.

Table 2.6 classifies Italian provinces by density of the network. Again, Rome stands out against the other provinces. Latina and Trieste (in the north-east) follow in second and third place. The table depicts a picture a bit different from the one described by Table 2.5. When looking at the network in relative terms, Northern Italy fulfils some importance in the Roman empire; nevertheless, the centre and the south hold their character of being at the core of the main street junctions.

Table 2.6 Roman road density (kilometres of road per 100 square kilometres of land area) by Italian province

Italian Provinces	All Roman roads density: km per 100 sq. km of land area	All Roman roads density: rank	Certain Roman roads: km per 100 sq. km of land area	Certain Roman roads density: rank	Italian Provinces	All Roman roads density: km per 100 sq. km of land area	All Roman roads density: rank	Certain Roman roads: km per 100 sq. km of land area	Certain Roman roads density: rank
Roma	46.57	1	20.97	1	Monza	5.69	56	0.00	107
Latina	29.74	2	11.70	3	Rimini	5.62	57	3.32	35
Trieste	20.93	3	12.73	2	Cagliari	5.59	58	0.00	86
Caserta	13.43	4	6.28	7	Udine	5.53	59	2.57	46
Lodi	12.74	5	1.03	78	Torino	5.37	60	3.08	38
Frosinone	11.90	6	2.81	44	Palermo	5.28	61	1.92	57
Olbia-Tempio	11.13	7	0.00	86	Messina	5.13	62	5.13	11
Napoli	10.92	8	5.12	12	Mantova	4.97	63	2.16	52
Lecco	10.87	9	3.95	22	Bologna	4.97	64	2.97	41
Ragusa	9.97	10	0.27	82	Catania	4.88	65	2.32	48
Brindisi	9.48	11	6.60	6	Brescia	4.77	66	2.00	54
Milano	9.38	12	4.67	15	Massa Carrara	4.68	67	0.00	86
Livorno	9.19	13	0.00	86	Nuoro	4.63	68	0.00	86
Cremona	9.12	14	4.19	18	Rieti	4.63	69	0.06	85
Gorizia	8.87	15	8.87	4	Ogliastra	4.61	70	0.00	86
Barletta-Andria-Trani	8.86	16	6.99	5	Alessandria	4.61	71	3.55	29
Avellino	8.73	17	1.57	63	Caltanissetta	4.58	72	0.00	86
Salerno	8.61	18	3.41	34	Medio Campidano	4.54	73	0.00	86
Potenza	8.49	19	1.80	59	Pescara	4.39	74	0.00	86
Siracusa	8.28	20	4.00	19	Vercelli	4.39	75	1.98	56
Oristano	8.23	21	0.00	101	Chieti	4.20	76	0.00	86
Taranto	8.21	22	3.93	23	Rovigo	4.20	77	1.20	72
Foggia	8.13	23	5.66	9	Pesaro-Urbino	3.99	78	3.99	20
Padova	8.05	24	1.71	60	Modena	3.98	79	1.04	77
Treviso	8.00	25	4.71	14	Asti	3.90	80	3.42	33
Cosenza	7.91	26	3.71	26	Novara	3.85	81	1.87	58
Reggio Calabria	7.82	27	5.76	8	Perugia	3.76	82	2.44	47
Como	7.78	28	5.38	10	Enna	3.61	83	0.36	80
Agrigento	7.69	29	0.19	83	Bolzano	3.60	84	2.27	50
Trapani	7.60	30	1.66	61	Aosta	3.45	85	3.45	31
Lecce	7.52	31	1.14	74	Pistoia	3.36	86	0.00	86
Viterbo	7.51	32	3.15	36	Fermo	3.10	87	3.10	37
Matera	7.49	33	1.53	64	Ancona	2.99	88	2.99	39
La Spezia	7.42	34	0.00	86	Imperia	2.99	89	2.99	40
Benevento	7.25	35	2.95	42	Ferrara	2.95	90	0.00	86
Bari	7.24	36	4.63	16	Parma	2.94	91	1.42	68
Catanzaro	7.24	37	3.82	25	Forlì-Cesena	2.92	92	2.17	51
Terni	7.19	38	3.65	28	Cuneo	2.90	93	1.05	76
Venezia	7.11	39	1.16	73	Bergamo	2.78	94	1.45	66
Teramo	6.87	40	1.48	65	Reggio Emilia	2.73	95	2.72	45
Campobasso	6.82	41	1.23	71	Prato	2.70	96	0.00	86
L'Aquila	6.77	42	3.70	27	Lucca	2.68	97	0.00	89
Crotone	6.75	43	2.82	43	Pisa	2.68	98	0.00	86
Ravenna	6.66	44	3.55	30	Grosseto	2.44	99	0.09	84
Verona	6.65	45	3.89	24	Varese	2.28	100	2.28	49
Firenze	6.53	46	1.11	75	Trento	2.23	101	1.43	67
Savona	6.49	47	3.42	32	Arezzo	2.19	102	0.00	86
Pavia	6.47	48	3.97	21	Vicenza	1.89	103	1.31	70
Piacenza	6.42	49	4.83	13	Sondrio	1.76	104	0.33	81
Ascoli Piceno	6.30	50	1.60	62	Belluno	1.55	105	1.31	69
Sassari	6.08	51	0.00	86	Biella	0.78	106	0.00	86
Vibo Valentia	6.05	52	2.04	53	Macerata	0.72	107	0.72	79
Isernia	5.87	53	2.00	55	Siena	0.48	108	0.00	86
Carbonia-Iglesias	5.72	54	0.00	86	Pordenone	0.00	109	0.00	86
Genova	5.72	55	4.27	17	Verbano-Cusio-Ossola	0.00	109	0.00	86

Source: Author's elaboration from McCormick, M. et al. 2013. "Roman Road Network (version 2008)", DARMC Scholarly Data Series 2013-5, and from Istat (2011)

2.9 Concluding remarks

In the late nineties-early two thousands the economic literature has grown interest in documenting the links between historical events and current economic and political outcomes. The evidence provided came along with the same result: history matters and has long-term effects on actual economic development. What has changed across the years is the channel through which history performs. First studies focused mainly on past colonisations, persistent formal institutions and factor endowments. More recent works started to concentrate on other channels through which history transmits its effects on the present economy, such as informal institutions, knowledge, technology, transport infrastructure. Alternatively to history, geography has been proposed as being affected by persistence and determining actual economic outcomes. The literature has been divided between one side suggesting that geography has a direct effect on current economic results and one side that refers to an indirect effect of geography through its influence on past events.

In this chapter the aim was to give account of the traditional and some more recent works, that retrieve in history an important factor explaining present economic results, and to present a new measure of Roman roads for its subsequently use in Chapter3.

The analysis of the literature was mainly devoted to the historical fact around which the new measure is constructed and has been split in four parts. This organisation was not trivial. First of all, the works that marked the birth of the literature, some innovative and very recent papers and other contributions that, instead, do not find a similar persistent effect have been considered. In a second step, the focus has been devoted to Italy, examining those works that have found in Italian history a long-lasting impact on modern outcomes. Then, it has been given account of those works which look at historical infrastructure. Lastly, the 'Roman world' and its persistence have been analysed using the recent contributions about this specific historical period. The latter part has highlighted how the economic literature has started, especially in last years, to put interest in the lasting effects of the Roman domination in modern world. The evidence coming from these works suggests that the impact of the Roman domination lasted not only in those centuries after the fall of the empire, but that it has been beneficial for the current development. Most of the research on the Roman world emphasises how development and urbanisation took advantage from the Roman domination, whereas less attention has been dedicated to the effect of the Roman road network. In 2015 two new works have tried to fill this gap, and with next Chapter 3 of this thesis, they are the most significant contributions in economic literature.

The Roman road network represents, in the view of historians and common people, one of the main traits of the Roman domination and the major sign of the glorious past of the Roman empire. The fascinating Roman world interested researchers as well as individuals in ancient times as today and the discovery of new details about how the Romans used to live and how were they able to establish a so magnificent empire is still subject of interest. Thanks to the availability of geo-coded data on the distribution of the Roman roads, it has been possible to create an original measure of Roman roads and to track the extension of the network in Italy. The new measure constructed produces three important contributions in economics. First, it adds a novel index of Roman roads for the Italian provinces to the existing group of old historical infrastructure measures. Second, it contributes to the literature on the persistent effect of history, allowing new empirical studies and applications. Third, it provides a method to generate additional measures of the Roman infrastructure for other countries and other territorial disaggregations. In conclusion, the construction of a new historical measure creates new knowledge, since it allows strongly related studies regarding the particular historical period which it is constructed for, as well as it contributes to the research of processes of long-run economic growth past and present.

APPENDIX 2.1 - THE ROMAN ROAD MEASURE: TECHNICAL NOTES

The strength and the potential of the Roman road measure described in Paragraph 2.8 and constructed for the subsequent empirical application in Chapter 3 lies in two main points. On the one hand, the availability of a new index useable for diverse investigations contributes to the existing data sets on the features and traits of the Roman world, enabling inferential analyses. On the other hand, it gives a method that can be simply followed to produce the Roman road measure in kilometres at different territorial levels and for one or more countries.

Whereas Paragraph 2.8 had the goal to present a brief descriptive on the newly created Roman road measure, this appendix aims to analyse more in detail those technical aspects linked to the use of GIS approaches.[21]

All elaborations have been computed using QGIS, a free open-source GIS software,[22] and STATA. Two starting layers have been used to compute the Roman road measure: the linear shape file of the Roman road network by McCormick et al. (2013) and the polygonal shape file of 110 Italian provinces provided by Istat. Both shape files need to be projected using the same spatial reference. The selection of the appropriate coordinate reference system (CRS) should consider the spatial reference system in use in the country or in the specific area the analysis looks at. Istat makes available the shape files of the Italian administrative borders (regions, provinces, municipalities) for the two main reference systems in Italy: UTM ED 50 Zone 32N and UTM WGS 84 Zone 32N. These reference systems are employed to homogenise the small and middle scale cartography at the European level. Important step of these starting phase lies in the choice of the reference system and in the attribution of the same selected CRS to both shape files.

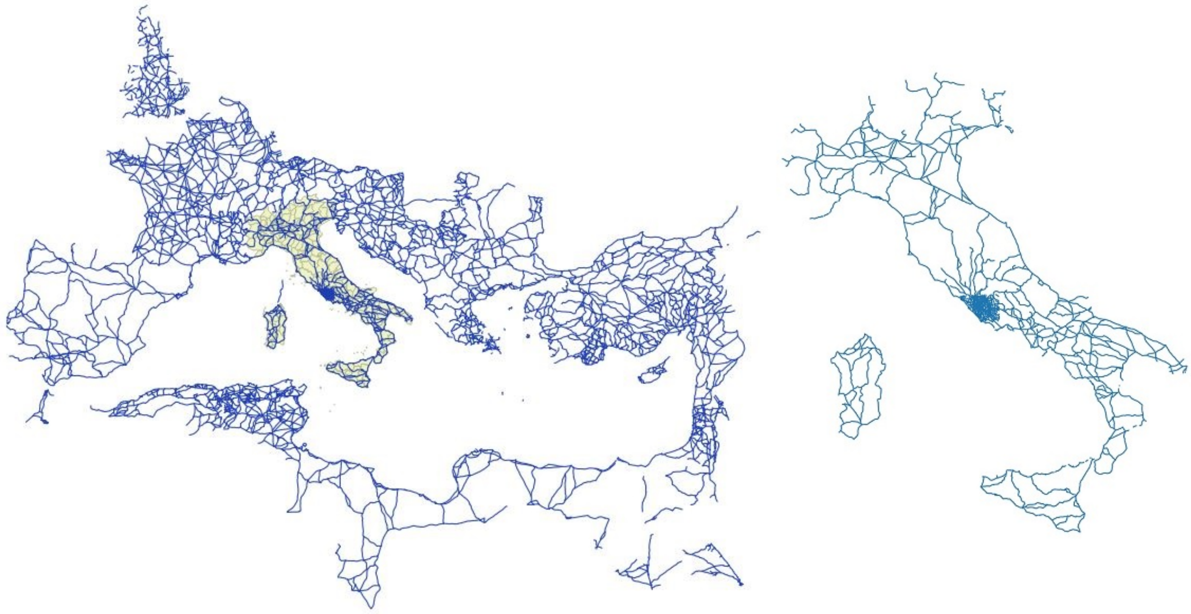
The two shape files with the same CRS serve the construction of a linear layer of the Roman road network for the sole Italian territory and the calculation of the Roman road network by Italian province.

The Italian linear layer of the Roman road network is derived superimposing the complete linear layer by McCormick et al. (2013) on the polygonal one with the provincial decomposition, as shown in Figure 2.6 (left), and using the latter as mould to derive the ramification of Roman road for the sole Italian territory (Figure 2.6, right). In this way,

[21] The instructions and technical notes of this appendix are the ones employed by the author of the thesis. The procedure described is not the only one available. Expert users of GIS methods might come out with the same result using different tools or following a diverse method.

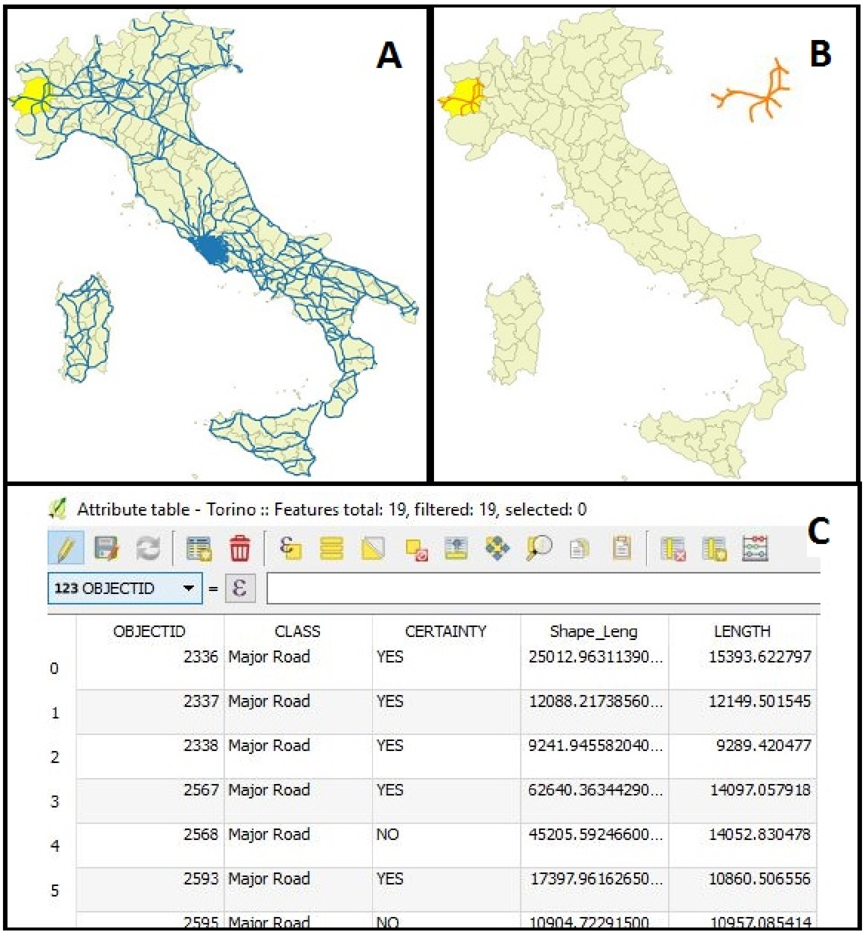
[22] Elaborations can be performed using ArcGIS as well.

Figure 2.6 Creation of the Italian linear layer of Roman roads



Source: Author's elaboration from McCormick, M. et al. 2013. "Roman Road Network (version 2008)", DARMC Scholarly Data Series 2013-5 and from Istat (2011)

Figure 2.7 Construction of the Roman road measure



Source: Author's draw

segments that are in common between Italy and adjacent countries are split and only the parts of the road included within the Italian territory are ascribed to Italy.

The computation of the Roman road measure in kilometres by province is more time consuming. It requires, as a first step, the construction of linear layers of Roman roads at the provincial level. 108 layers: one for each Italian province having Roman roads. Each linear layer of Roman roads is extracted one at a time. By selecting the boundaries of the given province on the polygonal layer of the Italian provinces, the linear layer of Roman roads (the complete one by McCormick et al. or the newly created linear layer for the sole Italian territory)[23] is projected and superimposed on the polygonal one. The borders of the selected province act as a cutter for the segments of roads that are in common between adjacent provinces, and only the parts of roads within the boundaries of the selected province are assigned to the province. This procedure is repeated 108 times. Each provincial layer of Roman roads includes the segments comprised in the given province. Although the starting linear layer by McCormick at al. (2013) provides the length of each segment, after having isolated and ascribed the segments to each provincial territory, the length in metres is not more valid since it refers to the complete section. Therefore, it is necessary to compute the extent of each segment within each provincial linear layer, repeating the process 108 times. Figure 2.7 displays and summarises the procedure in three boxes. Part A shows the first step of the calculation: the area of the province is selected (yellow area). In B the Roman roads within the boundaries of the selected province have been identified. In part C each segment of Roman road (identified by an id number) is linked to the corresponding length in metres computed using the specific tool.

The long procedure reported above, exploiting specific geoprocessing and geometry tools of QGIS, provides a measure in metres for each segment included within the provincial boundaries. A further step consists in collapsing and adding up by province all lengths in order to obtain a single measure for each province. Lastly, all measures are divided by 1000 to express the length in kilometres.

[23] It does not matter which linear layer of Roman roads is adopted (the complete or the Italian one) since the computation of the measure by province requires that the linear shape file includes only the roads within the provincial boundaries.

Chapter 3

The Long-Term Effects of the Historical Roman Road Network: Trade Costs of Italian Provinces

3.1 Introduction

Since school, teachers never grew tired of repeating that studying history is important. It is important because history allows to understand the present and because from history people should learn how to not repeat the same mistakes of the past.

The allure of the importance of history for economics started with the work of North in 1981; but it is only during the first two thousands that the increasing interest in economic research marks the birth of a new strand of economic literature. Economists started to move from the study of nearby determinants to the exploration of more remote and primordial factors, deeply rooted in history. Past dominations totally shaped the culture, the mindset and the institutions of peoples and territories, profoundly affecting the course of events in ancient times as well as nowadays. In fact, the persistent character of history makes culture, mindset and institutions a heritage that continues until today and that strongly determines current economic growth and development.

Integrating into this recent literature on the persistent effect of history and historical institutions on present outcomes, this chapter presents a novel empirical investigation on the long-term impact of history on modern economy, looking at that particular epoch that corresponds with the glorious Roman domination and the construction of the most important and long-lasting infrastructure in history: the Roman road network.

The main goal is to examine the impact of the existence of old Roman roads on current trade costs, assessing whether a stronger presence of ancient Roman roads is associated to a contemporary higher propensity to trade more with foreign countries than domestically. On the one hand, this chapter aims to validate a lasting impact of history on current economic results. On the other hand, it seeks to investigate whether the old Roman infrastructure left a mark on the present infrastructure. Along these lines, the chapter presents one of the first investigations on the effect of the Roman empire and its network system.

The Roman road network began to extend simultaneously with the military conquests of the Roman army. Starting from the city of Rome, the enlargement covered six centuries and three continents (Europe, Africa, Asia), coming to extend over almost forty countries at the peak of the domain. The first strategic Roman road was the Via Appia, constructed in 312 B.C. from Rome to Capua (close to Naples). In Italy, the system of Roman roads broadens in the entire peninsula, including the two main islands (Sicily and Sardinia). If considering a regional disaggregation (NUTS2), the Roman roads touched every region in Italy; looking at the NUTS3 disaggregation, 108 out of 110 Italian provinces have Roman roads.

The focus on Italy is aimed at the cause of identification. Roman Italy was constituted to create unity. Laurence (1999) depicts Italy as a whole composed by a series of cities connected by the road system. In this perspective, the road network is a structure between places, which connects them to create an artificial unity and sums up the fluidity of the regions of the Italian peninsula under Rome. According to this, Italy represents an ideal case study for two main reasons. First of all, the Roman domination has touched the whole Italian territory, shaping in a very strong and deep way its economy, society and space. The second reason can be found in the dual nature of Italy: a high-developed North-Centre and a less-developed South. Both traits are crucial in understanding whether and how much the concentrated existence of ancient Roman routes have affected trade costs today and, therefore, a stronger propensity to export and import abroad. From this perspective, Italy is perfectly suitable for examining how different past local institutions, like the Roman roads project, determine current outcomes with a persistent effect.

The selection of Italian provincial level data is not trivial. The use of NUTS3 provinces allows, on the one hand, to deal with a possible endogeneity problem between economic outcomes and historical transportation infrastructure, since it enables to gain more degrees of freedom from the local variability. On the other hand, as highlighted by Di Liberto and Sideri (2015), the provincial level in Italy represents a good geographical disaggregation to measure differences across provinces. Despite the central government has the main influence in determining institutions, the same institution seems to work differently, suggesting that some specific local factors affect the institution functioning.

The current measure of economic outcome used in the chapter is the indirect measure of trade costs, computed according to the top-down approach proposed by Novy (2013) for 107 out of 110 Italian provinces. The use of trade costs represents itself a novelty for two main reasons. First of all, contributions on the persistence of history exploit principally measures on development, like GDP, nightlight luminosity or population. This chapter goes against the grain, betting on a new and different index: the indirect measure of trade costs at the provincial level. Second and most important, this empirical investigation is the first, to the author's knowledge, that adopts a measure of trade costs not at the national level but at a sub-country level.[1]

As highlighted in Chapter 1, the indirect measure is thought to capture trade costs between country trading pairs (country i and country j); in this chapter, instead, the dyad is represented by the Italian province i and the country j. Therefore, trade costs measure the costs of trade between Italian provinces and world countries.

The empirical strategy has followed two different approaches. On the one hand, the empirical model has been thought to assess the pure and comprehensive effect of the Roman road network on current provincial trade costs. In order to capture this variability among the Roman road intensity measure and in order to exploit a larger sample, cross-section as well as panel data have been used. In particular, the Roman road measure is a time-invariant variable;[2] provincial trade costs, instead, have been measured for a 8 years period (2003-2010). Moreover, to control for historical, infrastructural and productivity differences across provinces, ad hoc variables have been included. Finally, fixed effects to control for unobservable and specific regional characteristics have been added to complete the model. On the other hand, the old Roman road infrastructure is adopted as an instrument for the current road infrastructure to evaluate whether provinces with a more intricate road network are more prone to trade internationally than domestically, since the influence of the old Roman road system on the current road infrastructure. Here the potential endogeneity problem from the use of the measure of current roads is avoided by the use of the Roman road measure, for which reverse causality issues are less likely to be the case.

The spirit behind the issue of the persistence of the Roman road network lies more in the attitude and mentality that the Roman roads built across the centuries rather than in the mere benefit of the road infrastructure. In fact, it must be considered that the development of land

---

[1] Paragraph 1.5 in Chapter 1 and following Subparagraph 3.4.1discuss about the use of the indirect measure of trade costs at the sub-country level.

[2] It refers to 117 A.D.

transportation was more related to the cultural change from a society based on the city state, as the ancient Greece was, to a state that was not a 'nation' but associated with a dispersed citizenship, like the Roman empire. During the Roman domination, Italy can be seen as a mixture of peoples across space, rather than a specific territory with one specific community. This led to a change in the mentality of space and time, revolutionizing the importance in considering the territory and the time (Laurence, 1999). Although moving around was expensive and before leaving on a journey people were superstitious, the Romans enjoyed travelling for business and for amusement, because mobility was a keynote of the society (Chevallier, 1976). Accordingly, von Hagen (1967) argues that, through its roadways, Rome was able to run a systematic control of the world, making the Romans a mobile civilisation. The fundamental idea behind this chapter lies exactly on this concept of openness and connectedness: history, the ancient Roman civilization, through the road infrastructure, could have had a positive long-term effect in terms of more propensity to trade with foreign peoples, and this effect could have lasted for two thousand years. Hence, the positive impact of the Roman empire, in the view of this chapter is more 'mental' than 'physical' (i.e. modern roads built on ancient Roman roads). Economists argue that, although the transport costs of that time, trade across the extended Roman empire during the first two centuries A.D. was possible; and it was possible thanks to a unified political system that could effectively enforce rules. In other words, it is not the transport cost but the cost of transacting that is the key obstacle that prevents economies and societies from realising well-being (North, 1989).

The beneficial ʽmental openness' effect played by the Roman infrastructure is well explained by Laurence and Trifilò (2015, p. 110): "*For the Roman empire, we contend that there is a sense by which places on a named highway or at a coastal location had a greater level of connectivity than those at a distance from such routeways across the empire*". However, settlements close to roads, rather than ports, had a greater impact in terms of monument building and infrastructural investments by the Roman state. The purpose was to make travellers familiar with a set of buildings and images typical of the ʽRomanness'. In the view of Laurence and Trifilò (2015, p. 111) "*This implies that land transport was a greater facilitator of cultural integration than the phenomenon of sea transportation*". Accordingly, Jennings (2011) finds that during Roman times the ʽtime-space compression', i.e. the perception that the world is smaller and distances are shorter, originated from the new opportunities of inter-regional trade and traded commodities. And these achievements have been possible thanks to the use of the donkey for the traction of goods and the creation of a consistent road system. Moreover, Laurence and Trifilò argue that locations along the roads were the places where the 'time-space compression' occurred, since the predictability of the land travel compared to the uncertainty of the sea travel. The focus on the road infrastructure and on the resulting land transport, rather on the river or sea transportation, and the development of a measure of Roman roads in kilometres, rather than relying on the information on ports or navigable rivers during Roman times, come exactly from the key role played by the road system in creating an interconnected world. Although rivers and sea had a fundamental role in creating openness, connectivity and mobility, the distinctive character of the road system lies in the potential the road had: the development of cities and the settlement along the roads, the monuments construction, the improvement of the road surface to make journeys faster, the creation of the Antonine Itineraries, a map register of distances and stations along the main roads, are all symbols of the constitution of a global Roman culture based on mobility and connectedness.

This chapter is based on seven paragraphs. Paragraph 3.2 discusses the typical reverse causality problem that may rise when investigating the effects of infrastructure on economic outcomes and how this seems to be less the case for the Roman road network. Three distinct motivations are provided in supporting this claim. In paragraph 3.3 a preliminary and pioneering investigation of the relationship between Roman roads and geography is provided. Here the aim is to better understand how the Romans built their roads in topographical terms: whether some particular territories were avoided, whether they circumvented specific geographical barriers or whether they went beyond physical obstacles in order to get roads as straight as possible. Paragraph 3.4 has a twofold structure. It is devoted to describe data and to provide figures about Italian provincial trade costs, to better understand what are the patterns of the dependent variable and to highlight whether the provincial level represents a good disaggregation in terms of variability. Moreover, it presents data and descriptives on the accessory independent variables that complete the analysis (current roads, historical controls and productivity index). Since the extensive and complete focus on the new Roman road data set in Chapter 2 and since the descriptive analysis on Roman roads provided in paragraph 2.8, Paragraph 3.4 does not present any focus on the Roman network. The two main empirical approaches adopted to investigate the effects of Roman roads are presented in Paragraph 3.5, discussing the motivations behind their choice. Paragraph 3.6 discusses the results coming from the inference analysis and from robustness checks. Finally, Paragraph 3.7 provides some concluding remarks. The chapter ends with Appendix 3.1, where are addressed all inescapable post-estimation tests for the IV approach exploited, and Appendix 3.2, which includes all tables of the robustness checks.

3.2 A possible endogeneity problem?

The investigation on the (beneficial) effects of infrastructure projects is not plain, since the inference analysis is plagued by some problems that can mislead the researcher. Economic and econometric literature identify the main drawback in the endogeneity problem. As highlighted by Wooldridge (2002), the endogeneity arises from three main sources: measurement error, omitted variables and reverse causality. The more serious cause among the three is represented, in the case of infrastructure, by the reverse causality, according to which the direction of causality is from development to infrastructure rather than from infrastructure to development. Higher investment in infrastructure and higher economic output do not necessarily reflect the fact that infrastructure is boosting economic development, but it might indicate that in wealthier areas there is an increasing demand for infrastructure or that new infrastructure projects are placed in those regions with a higher income.[3]

The endogeneity drawback behind the use of infrastructure in empirical models has been highly discussed in literature,[4] in conjunction with the econometric solutions that help to deal with this problem. However, as highlighted by Brooks and Hummels (2009), there is a significant time between the plan of the road and the moment when the road is effectively constructed and employed; on this perspective, the road infrastructure can be seen as an exogenous variable.

It might be wondered if this is also the case with the old infrastructure. Donaldson (2015) argues that the effect of historical transportation infrastructure is characterised by a potential simultaneity problem: roads and railways are often constructed to connect already active trade regions, while trade relationships between regions often emerge after the construction of infrastructures or road improvements. On this view, it can be argued that the relation between the existence of Roman roads and a higher propensity to trade today might suffer of an endogeneity problem. Roman roads might have been constructed to reach economic prosperous and flourishing territories, and these promising conditions lasted until today, reflecting more active trade provinces. Looking at Roman empire, this is less likely to be the case. It is more likely, instead, that the construction of the road represented a benefit for the territory, allowing the development and sustainable growth in Roman times as well as across the centuries. On this point, Chevallier (1976, p. 116) points out how "*As a rule, earlier sites*

[3] McCulloch (2009) underlines that, if richer areas attract new investment in infrastructure and poorer areas do not, then the direction of causality between income and infrastructure is hard to determine.

[4] See Paragraph 3.5 for a more specific analysis.

were avoided by Roman roads, especially the great Imperial highways, which were unconcerned with local interests and small settlements. [...] The road often attracted the village, but when the ancient road itineraries name a civitas, it does not mean that the route went through the town itself: occasionally it simply skirts its territory". Moreover, Bosker et al. (2013) confirm that Roman roads get rid of reverse causality issues, since they favoured the subsequent expansion of urban centres in those territories where roads passed through, rather than being constructed for already existing settlements.

In order to definitively exclude endogeneity problems coming from simultaneity concerns, the research effort of this chapter consists in understanding why and how Roman roads were built. The purpose and the methods of the construction of the roads are derived from historical sources and they represent the two reasons that support the absence of reverse causality issues: Roman roads were constructed for military purpose and they were straight. The military and the engineering reasons help to understand how, behind the decision to construct roads, the mere economic purpose is not at the basis of the motivations. The process followed by the Romans to build roads was highly systematic and methodical, establishing well-known regularities. Moreover, as a further support of the absence of an endogeneity problem, the construction of the most important Roman road, the Via Appia, represents a perfect case study to illustrate how the direction of causality is from the infrastructure to the economic outcome.

In Subparagraph 3.2.1, the military reason is presented exploiting sentences and statements from historical sources.[5] Subparagraph 3.2.2 explains the engineering reason according to which the Romans built their roads as straight lines. The practical case of the Via Appia is described in Subparagraph 3.2.3.

3.2.1 The military reason

The purpose to build roads was always militaristic: they were constructed for war campaigns. Countless historical works confirm that the original function of Roman roads was designed to deal with the military expansion of the empire.

"*The public road-system of the Romans was thoroughly military in its aims and spirit: it was designed to unite and consolidate the conquests of the Roman people, whether within or without the limits of Italy proper*" (Dictionary of Greek and Roman Antiquities - Smith, 1890). Even after construction, they had no significant economic impact, since the cheaper modes of transport during that time were by river or sea (Finley, 1973). Laurence (1999)

[5] A more colloquial presentation of the military reasons is provided in Paragraph 2.6 of Chapter 2.

clearly explains how roads were planned and designed in order to provide troops with the essential means, to guarantee an efficient repositioning and to easily facilitate the army movement roads were straight with intermediate safe stops.

The military reason is also strongly supported from the Latin literature. "*After having pacified Liguria, Aemelius had his army build a road from Piacenza to Rimini to join the via Flaminia*" (Livy, 59 B.C.-17 A.D.). In his "Encyclopedia of antiquities, and elements of archaeology, classical and medieval" of 1843, Fosbroke reports how the Anglo-Saxon ancestors named the Roman roads 'military ways' and how they thought that the construction of small roads had a more military utility than that of large ones. Chevallier emphasises the importance of the army's role in the case of the main roads. He remarks that "*[...] the majority of main roads were pioneered by military operations. For example, on its return from the first Samnite war (343-40), the Roman army did not come back along the via Latina, but followed the coast through the territory of Aurunci, thus blazing the trail of the Appia on a line that had already been known to traders, at least since the hegemony of Etruria. In the early third century, operations against the Umbrians of Mevania and Narnia and against the Senones took into account the route that became the Flaminia. Great strategic road were built by the military in Gaul under Agrippa from BC 16-13 in Dalmatia and Pannonia under Tiberius from AD 6-9, in the Rhineland and the Danube valley under Claudius, and in Asia Minor under Flavians*" (Chevallier, 1976, p. 85). During the fourth, third and second centuries B.C., Laurence (1999) underlines how the Roman road network played a central role in guaranteeing the Roman supremacy in the world. The Roman road network development and the foundation of colonies had the same purpose: to establish Roman military control and to avoid revolts in new conquered territories. On this point, Rosenstein and Morstein-Marx (2006, p. 608) point out that "*Hand in hand with the establishment of colonies went the construction and the extension of the Roman road network. Sometimes following preexisting routes and sometimes adopting new ones, the roads had an overtly military purpose - in this case to allow Roman armies to travel swiftly Italy and to provide links with the colonies. [...] The building of the Via Valeria in 307, extending the Via Tiburtina eastward into the Apennines, likewise appears to be the precursor of the military campaigns against the Aequi in 304 and the foundation of the colonies at Alba Fucens and Carseoli*".

3.2.2 The engineering reason

The engineering mark of the Roman network is the second mainstay against reverse causality. The best-known feature of Roman roads was their straightness: the Romans ran straight lines between two strategic points and built the road as segment connecting them. In fact, Cornell and Matthews (1982) point out how the first step of the road construction consisted in marking the path of the road with stakes and furrows in a way that was as straight as possible, employing line of sight measuring tools. The Romans preferred direct and straight roads, because with this outline it was easier to circumvent obstacles, to avoid ambuscades and to keep away from population settlements. Moreover, straight roads were easier to secure (Gleason, 2013). As suggested by Poulter (2010) and as remarked by Bishop (2014), the Romans often did survey on the points from which starting and ending the road, that were typically placed on the top of a hill, and the road came along as a segment connecting them. Von Hagen (1967), on the constitution of a mobile civilisation throughout the continent, argues how this has been possible thanks to well-engineered and straight line roads.

The straightness of the road and, more specifically, lines connecting two points or main cities are the focal point of the identification strategy of a strand of economic literature that started with the work of Banerjee et al. (2012)[6] and continued with the main contributions of Atack et al. (2010) and Faber (2014). These contributions, although they might appear faraway from the majesty of the Roman world and from that literature on the persistent effects of the Roman domination, are very close to the insights behind the construction of the Roman road network. Moreover, the work of Banerjee et al. (2012) provides further validation to the research question of this chapter, since it is focused, as well, on the long-term effects. Indeed, Banerjee et al. (2012), as discussed in Chapter 2, are aimed to investigate whether those areas with better access to transportation networks have better economic outcomes in the long run. On this aim, they find in straight lines a proxy for transportation infrastructure and an optimal instrument to assess the access to infrastructure. To examine this issue, they look at the Chinese infrastructure. Chinese infrastructure represents an ideal case study, since, in the late nineteenth and early twentieth century, China started the construction of railroads which linked the historical Chinese cities and strategically military ports of China to each other. The line, connecting the cities and the ports, is the proxy of the transportation infrastructure and it is used to disentangle the areas which benefited from the infrastructure, due their closeness to the line (treated areas), from the areas which could not take advantage from the infrastructure, since their distance (non-treated areas). On these basis, the empirical strategy examines the correlation between the distance to the nearest straight line and the performance during a current twenty-year period of rapid growth. According to the authors, the identification

[6] The earlier version of the paper refers to 2004.

strategy, based on identifying the infrastructure network with straight lines, is an exogenous source of variation in access to transportation and addresses the problem of endogeneity, since it exploits the fact that these networks tend to connect historical cities. The work of Banerjee et al. (2012) is particularly interesting and motivating. In fact, in order to avoid a possible endogeneity problem using the current infrastructure, they exploited the historical Chinese infrastructure. And this historical infrastructure built on straight lines allowed the authors to deal with reverse causality issues as well as to rely on the straight-lines approach, to identify treated and non-treated areas, to control for distance to the terminal cities and distance from the river.[7]

The work of Banerjee et al. (2012) provides a further proof on how the engineering reason of the Roman road network supports the absence of an endogeneity problem, on how the straightness of the roads seems to be the rule of the old historical infrastructures and on how it enables identification strategies of being robust to endogeneity issues.

3.2.3 The case of the Via Appia and the conquest of Greece

The Via Appia was the first large strategic road that radiated from the city of Rome to outside. It connected Rome to Brundisium (modern Brindisi, Apulia), in South-East Italy, and was the first demonstration of the military power, the leadership and the engineering abilities of the Romans. Started in 312 B.C. on the Roman censor Appius Claudius Caecus' request, the road had the tactical purpose to allow the movement of the army outside the region of Rome during the Samnite Wars.[8] It was constructed in subsequent lots, following the progress of military campaigns, and the completion of the road was achieved in 191 B.C., when it reached Brindisi.

This paragraph has the precise intent to demonstrate, how the construction of the Via Appia across the centuries represents a clear and real example of how it served for military purposes and how the final aim to reach some strategic territories resulted in a long road that went through areas completely uninteresting to the Romans, but that, nonetheless, benefited from the presence of the road. The case of the Via Appia is not isolated, it represents rather the norm of Roman roads. The fact of being constructed for military campaigns and of reaching strategic points does not prevent that featureless territories took advantage of having a road nearby.

[7] Distance to the terminal cities and distance to the nearest navigable river are two caveats. See Banerjee et al. (2012) for detailed information.

[8] Chambers's Encyclopædia, Vol. 1, p.490.

To better understand the Via Appia construction, the simultaneous military conquest of the southern part of the Peninsula and the forthcoming expansion of the Roman empire in Greece, Figure 3.1, Figure 3.2 and Figure 3.3 have been drawn as a support for the motivations behind the constitution of the Roman road network and the absence of a potential endogeneity problem.

The construction of the Via Appia started in 312 B.C. During that time, the Roman empire was comprised, as shown in the upper part of Figure 3.1, in those territories belonging to the Latine League[9] and corresponding, today, with the provinces of Rome and Latina. Between 500 and 400 B.C. the Romans had already defeated the neighbouring peoples living in central Italy (the Etruscans, the Latins, the Sabines, the Lavinii, the Tusculi, the Aequi, the Volsci, the Aurunci and the Veientes), controlling a limited area in central Italy, and at the end of the fifth century B.C. the Italian peninsula was under the control of the Celts and the Gauls in the North, the Romans in Centre-West, the Samnites and the Greek colonies (Magna Graecia) in the South. It is exactly in this time, that the decision to build the first track of the Via Appia was taken and that the Romans started to show interest in the southern part of Italy. But, they were not the only ones. The Samnites were an italic people living in south-central Italy, in those territories that correspond today to North-East Campania, northern Apulia, Molise and southern Abruzzo, who conflicted with the Romans, due the military expansion politics of both peoples. At first, the Romans and the Samnites concluded a non-aggression pact, agreeing to expand their possessions throughout distinct directions, but this treaty was irremediably broken when the expansion directions of the two peoples converged. On the one hand, the Samnites aimed to extend their area in the western part of Campania. On the other hand, the Roman expansion goal was, first, to expand their territories in southern Italy, to obtain new lands for the growing Roman population and to entertain commercial relationships with the Greek merchants,[10] but, later, the final challenge was the conquest of the Magna Graecia and to extend the control over the Mediterranean Sea, where most of the trade occurred. The war with the Samnites began in 340 B.C.[11] and simultaneously the first segment of the Via Appia was built. Capua was the '*casus belli*' that lead to the First Samnite

[9] The Latine League is a term, coined by modern historians, that identifies a coalition of villages and tribes settled in central Italy, surrounding Rome and that had the primary role to guarantee the mutual protection against external enemies (Cornell, 1995; De Francisci, 1968).

[10] Musti (1990).

[11] The war with the Samnites lasted 53 years.

Figure 3.1 Roman empire and Via Appia in 312 B.C.

Source: Author's draw

War (340-341 B.C.). It was the third city in the peninsula, after Rome and Tarantum (modern Taranto, Apulia),[12] a flourishing centre of the Magna Graecia and a strategic hub, since its coastal position. For these reasons, the Samnites were highly interested to extend their control over it. Capua asked for the Romans' protection, and one year later, after the start of the First Samnite War, with the victory of Rome, Capua was annexed to the Roman territory and became part of the Roman empire. The earlier Via Appia linked Rome to Colli Albani,[13] but in 312 B.C. Appius Claudius Caecus decided to renovate the road and to extend it until Capua, as shown in the lower part of the Figure 3.1. Appius Claudius was one of the first to understand that the merger between the Roman and the Greek worlds might represent a beneficial opportunity for the Romans.[14]

Capua represented a strategic point during the First Samnite War, and, with the annexation to Rome, it favoured the alliance and the connection between Rome and the territories in Campania.[15] Moreover, the foundation of Cales (very close to Capua) as Roman colony was crucial during the second stage of the First Samnite War as well as during the Second Samnite War (326-304 B.C.). In fact, its favourable position, lead Cales to being the strategic zone where the war occurred and where most of the military battles took place. Moreover, Cales was placed along the Via Latina, constructed 22 years before the Via Appia. Although the favourable position of Cales, the Via Latina was not so strategic for the new Roman military campaign as the Via Appia was. Quilici (1989) clearly explains how, with the construction of the Via Appia, the censor Appius Claudius provided an alternative to the Via Latina, since the Appian way was constructed along the coast and, for this reason, represented a more backward and secure road for the military front. Accordingly, Laurence (1999) suggests how the Via Appia represented a strategic and secure answer to link Rome to its Campania allies and a demonstration of the Roman power.

In 312 B.C. the expansion of the Roman empire was only at the beginning. It took another 70 years to complete the expansion over the entire southern Italy. The fight with the Samnites ended in 290 B.C. with the victory of the Romans in all three Samnite Wars, extending their control in Campania. The Romans prepared to confront Pyrrus, the Greek King of Epirus. Rome pressed to conquer the Magna Graecia and to reach Taranto for its key strategic position as main port in the Mediterranean Sea. After a first defeat of the Romans, they got the better on Pyrrhus in 275 B.C., in the final battle of Maleventum (later Beneventum, now

[12] Quilici (1989).
[13] Colli Albani are a group of hills in the countryside of Rome.
[14] Clemente (1990).
[15] Quilici and Quilici Gigli (2004).

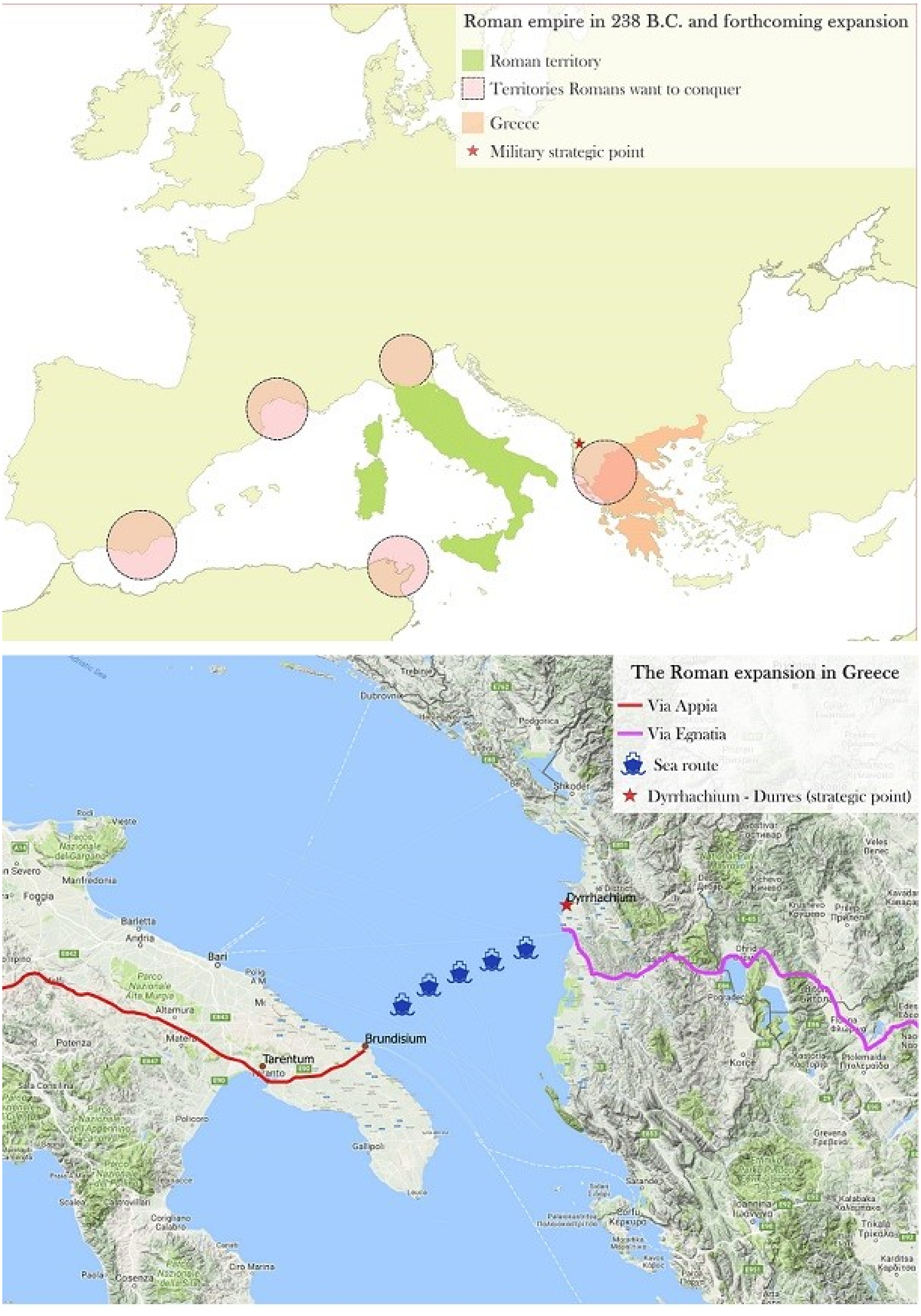
Figure 3.2 Roman empire and Via Appia in 238 B.C.

Source: Author's draw

Benevento, Campania). Meanwhile, the Via Appia was extended to Benevento first and to Taranto then, and the Romans secured most of southern Italy. The conquest of the south-eastern Italy was not complete, the Romans aimed to extend the control over Brindisi. Brindisi was an important seaport in the Mediterranean Sea that enabled to reach Greece and the Orient. In 238 B.C., the situation of the Roman supremacy was the one provided in the upper part of Figure 3.2. The Romans controlled the entire Centre and South of the Italian peninsula, including the three main Mediterranean islands (Sicily, Sardinia and Corsica), obtained after the Punic war of 264-241 B.C. (ended with the Carthaginian defeat). At the same time, the Via Appia (lower part of Figure 3.2) was extended, reaching Brindisi in South-East. During this period, as shown in the upper part of Figure 3.3, the Romans not only ruled the two third of the entire actual Italian territory, but they aimed to expand to North Italy (in those territories that were under the control of the Celts), to Gaul, to Spain, to North Africa and to Greece. In fact, as remarked by Gedacht (2004), the Mediterranean Sea was an important source for the enlargement of Roman empire and the Romans planned to conquer those territories that bordered the Mediterranean Sea. Greece, in the South-East, was one of the directions that the Romans followed to gain the access to the sea trade routes with the Orient. The conquest of the Greece showed in the lower part of Figure 3.3 helps to understand how the Via Appia extension until Brindisi was fundamental to arrive in Greece. Perri (2016) explains that the extension of the Appia until the port of Brindisi allowed it to represent the starting point from which the Romans could expand to the Orient. Brindisi' seaport had an important role indeed from the First Punic War (264-241 B.C.).[16] After the First Punic War, trade in the Adriatic Sea, the northernmost arm of the Mediterranean Sea, increased, but the continuous attacks on trading vessels of Rome's Italic allies by Illyrian pirates, led to the First Illyrian War (230-229 B.C.). With the defeat of the Illyrians, the Romans stopped the Illyrian seaborne assaults and gained control of the strategic ports of Epidamnos (changed in Dyrrachium, modern Durres, Albania), Apollonia (old Illyrian city in modern Albania) and Corcyra (modern Corfù), starting the conquest of Greece. Dyrrachium, Apollonia and Corcyra became Roman colonies and played an important role during the Second, Third and Fourth Illyrian Wars. Dyracchium was developed as major military and naval base and served the

[16] "*Brindisi, infatti, con l'estensione della Via Appia fino al suo porto, divenne la testa di ponte per l'espansione romana d'oltremare verso Oriente, e il suo porto giocò un ruolo importante fin dalla prima guerra punica [...] Nel 229 aC, durante la guerra contro il pirati illirici che con base in Albania infestavano tutto l'Adriatico, il porto di Brindisi ospitò l'intera flotta di 200 vascelli che, al comando del console Gneo Fulvio Centumalu, salpò verso Apollonia per sottomettere la regina Teuta, mentre l'esercito, con due legioni di 20.000 fanti e 2.000 cavalieri, guidato dall'altro console, Lucio Postumio Albino, era traghettato da Brindisi*" (Perri, 2016, p. 24)

Figure 3.3 Romans' expansionist objectives and the conquest of Greece



Source: Author's draw

Romans to easily reach the opposite coast of the Adriatic Sea.[17] As shown in the lower part of Figure 3.3, the construction of the Via Appia, and its enlargement step by step until Brindisi, favoured the use of Brindisi as starting seaport in the subsequent conquest of Epirus, throughout the landing at the Macedonian coast and thanks to ally ports in the opposite coastline, as Durres was. Moreover, starting from Durres,[18] the Romans were able to penetrate the Greek territory, thanks to the construction of the Via Egnatia, up to Byzantium (later Constantinople, now Istanbul). According to Hacquard (2003), the Romans constructed two roads to leave Italy, one of them (Via Domitia) to the west, to Gaul and Hispania, and the other one (Via Egnatia) to the east, to Greece and Asia Minor. Using the Via Egnatia, the Romans were able to expand throughout the Orient.

Using historical sources, this paragraph challenged to provide a descriptive and a practical overview of the expansion of the Romans to new lands and the simultaneous construction of roads, using the Via Appia as natural experiment to understand whether the Roman road network is affected by an endogeneity problem. Putting together lots of references, reports and historical narrations, the picture that emerges seems to disclose the absence of an endogeneity problem. Three facts come into sight from the narration above: 1) the instrumental role of the roads in the military conquest of new territories; 2) the development and the enlargement of roads occurred by strategic points: the Romans ran new segments using two tactical cities; 3) the construction of roads was a step by step process, but with a vision of the forthcoming expansion (Appius Claudius was far-seeing, expecting the future expansion to the Greek world). From these facts, it might be deducted that some territories, that were located along the direction to get to the strategic final point, were crossed by Roman roads although the Romans had no economic, military or tactical interest to cross them. In other terms, they benefited from the presence of the Roman road just because they were halfway between the initial and the ending point.

The meaning of the entire paragraph might be summarised by two sentences of Chevallier in his book "Roman roads" of 1976 (pp. 132-133): "*Rome first turned her attention to Campania. The via Appia, first major link with the south, was planned by the censor Appius Claudius Caecus in 312 B.C., originally as far as Capua. It was then extended (via Trajana) via Beneventum, Aeclanum, and Venusia, Horace's homeland, Tarentum to Brundisium (264), the embarkation port for Greece (in Epirus, the route was continued by the via Egnatia [...]*".

[17] Morgan (1980).

[18] Information on the starting point of the Via Egnatia are discordant. Some sources refer to Apollonia; the most refer to Dyrrachium.

3.3 Roman roads and geography

The relation between Roman roads and geography might provide important information on why Roman roads are located in given areas and not in others, on whether the Romans cared about geography and, if so, on how geography ruled the Roman road network.

Here the focus is on the topography of the territory. The knowledge of the connection between Roman roads and geography might help to better address the potential endogeneity problem, adding a further support against the existence of constraints in the use of Roman roads in empirical analyses. Accordingly, Ramcharan (2009) argues that topography can shape both the within-country spatial distribution of road infrastructure and the economic activity, and, if so, it represents a potential unobserved factor that is correlated with both road building and economic performance. Omitted variables are one of the main sources of endogeneity, and to control for them helps to avoid the drawbacks and weaknesses of the inferential analysis. Moreover, Ramcharan (2009, pp. 559-560) argues how "*[...]countries with rougher surfaces also have less dense surface transport networks: a 1% increase in roughness is associated with about a 1% decline in the number of kilometers of roadway within a country*". This evidence is consistent with the road construction literature,[19] according to which the shape and the structure of the territory heavily determine the cost and the time required to construct and to maintain roadways and railways.

The path dependency of the Roman roads is clearly explained by Lopez (1956, p. 17) who describes "*That the network of roads should be convenient and economic was none of their*[20] *business. That is why the Romans built narrow, precipitous roads along the mountain crests rather than the valley bottoms, sometimes driving straight for their goal over gradients of one in five*". On the same view, Margary (1973) describes that, in order to get roads as straight as possible, the Romans built roads with steep slopes or passing through mountainous territories. Bishop (2014), referring to Britain, finds in Hindle (1998) and in Welfare and Swan (1995) that straight long sections were typical of the major Roman roads. However, the straightness was typical also when roads accommodated to changes in the structure of the terrain. Most of the non-major Roman roads exhibit some deviations from the main path. These variations in the courses of the roads were typically short and, rather being curvy, they were subject to a change in the degree of the layout. This represents the typical mark which distinguishes the Roman infrastructure from the modern infrastructure. In light of this, Bishop refers to the Roman roads as '*surveyed roads*' which origin from a geometric-linear perspective in

[19] Among others, important contributions are those of Aw (1981) and Tsunokawa (1983).
[20] Lopez (1956) refers to the Romans.

conceiving the network. Current roads are, instead, in the words of Bishop, more linked to the '*line of desire*', since there is not a geometric outline behind the planning of the network, but rather a preference to follow the shape of the nature.

Roman roads were subject to deviations only when large obstacles impeded Romans to overcome them with engineering infrastructures, such as bridges. On this point, Bishops suggests that topographical and natural barriers were the main cause that led to adjust the course of the road in order to avoid them. Nevertheless, demanding territories were not avoided by the Romans. As pointed out by Richard (2010), the Romans, whenever possible, built road supports, like embankments or dykes, bridges, tunnels to cross hills, mountains, and marshlands. In Italy, their roads in the Alps and the Apennines had, according to Richard, steep slopes and can be defined as ancient motorways since they allowed the movement of pedestrians, horses and wagons.

Historical sources and more recent investigations agree on the presence of Roman roads in mountainous and steep areas. The Digital Atlas of Roman and Medieval Civilizations (DARMC) provides an interesting on line map source which allows to appreciate the mountainous passes used by the Romans.[21] This information comes from the "Barrington Atlas of the Greek and Roman World" (2000), and a focus on the Italian peninsula highlights a large number of passages on the Alps and the Apennines. The shape file on the Roman roads network analysed in Chapter 2 and the map of the mountainous passes are an indisputable sign on how the building abilities of the Romans overcame the geographical barriers. However, here the point is not to confirm or underline the existence of mountainous Roman roads, but it is rather to understand and explore whether more mountainous territories or areas with higher elevation are more prone to having less kilometres of Roman roads, as suggested by Ramcharan (2009) for current roads. This kind of analysis is not neutral with respect to the empirical application, that is at the core of this chapter. It should be considered that 35 percent of the Italian territory is represented by mountains, 42 percent is composed by hills, 23 percent is plain. It should be also taken into account that Italy was the first area where Roman roads were constructed and which experienced the primordial engineering techniques of the Romans, hence the focus on the Italian territory is particularly suitable.

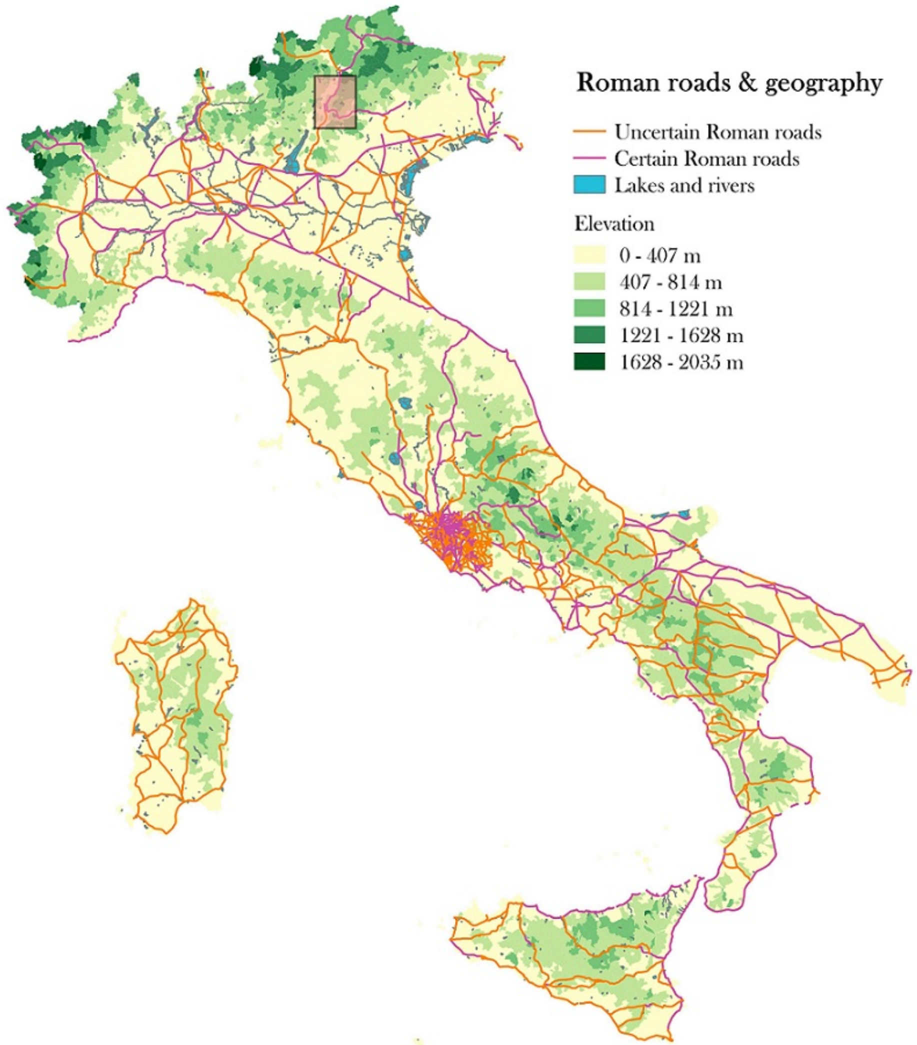
The investigation on the relationship between Roman roads and geography, where geography subtends the topographical features of the territory (percentage of mountainous territory and elevation), has been performed adopting two different approaches.

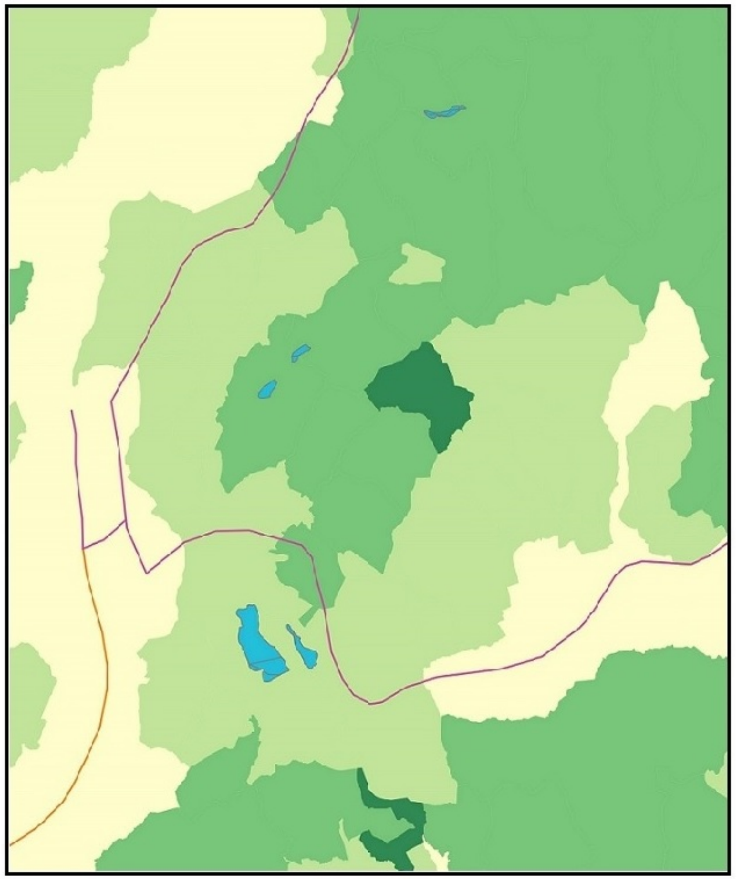
[21] See http://darmc.harvard.edu/maps for more information.

The first approach has consisted in identifying in which territories Roman roads were constructed. On this purpose, data on the elevation at the Italian municipality level provided by Istat have been geo-coded and mapped using the polygonal layer of the Italian borders. To perceive differences in altitudes, five different classes have been distinguished. The Italian map has been completed with the information on lakes and rivers using Corine Land Cover (CLC) data in order to control for watercourses and basins. In a second stage, the linear map of the Roman road network at the Italian level has been superimposed on the polygonal one and Roman roads have been distinguished between certain and uncertain. As remarked in chapter 2, the certainty refers to the direction followed by the road and this information is particularly significant, since the course of the road identifies where the road has been constructed. The graphical approach now described is displayed in Figure 3.4. The observation of the left part of Figure 3.4 immediately allows to notice that there is a quite homogeneous distribution of Roman roads among the diverse elevation territories. Darker areas, where the elevation is higher, are not disregarded by the Roman infrastructure. In the South-Centre, there is a high concentration of Roman roads in the Apennines, the second mountain range in Italy. Nevertheless, in the North the highest concentration of Roman roads is along the Po Valley, where, clearly, the average elevation is lower. In order to better appreciate the Romans' choice in building roads, the right part of Figure 3.4 zooms in a selected area of North-East Italy. The chosen area is not neutral, since it includes four different elevation zones, lakes and a trait of Roman road that is certain. It is possible to observe how the course of the selected Roman road passes through plain and more elevated areas. The road does not circumnavigate the lake, where the altitude is lower, but it crosses a more elevated area. However, the information taken from Figure 3.4, although important, should be taken cautiously. A first and preliminary conclusion is that Roman roads do not care about geography, but integrating the map with information on urbanisation and settlements during old Roman times would give a better and more complete view of the relationship between geography and Roman roads.

The second approach exploited to investigate the relationship between the existence of Roman roads and the geography of the land relies on the simple statistical information provided by correlations, in order to assess whether there is a dependence between the Roman infrastructure and the topography of the territory. Ramcharan (2009) argues that rougher territories have less kilometres of roadways. This view implies that in rougher territories roads are sparse. However, more precisely, it might be argued that rocky areas have less kilometres of roads since the difficulty to construct roads in demanding areas, but more kilometres of

Figure 3.4 Roman roads and geography: locationing by elevation area





Source: Author's draw from Istat data, Corine Land Cover data, McCormick, M. et al. 2013. "Roman Road Network (version 2008)", DARMC Scholarly Data Series 2013-5

roads are needed to reach those territories. In this view, the relationship between kilometres of roads and geography of the territory should be carefully assessed. On the one side, it should be considered which measure of Roman roads (simple or in density) is more suitable to use in the correlation analysis at the provincial level. The measure of Roman roads in density is potentially able to say something about the nature of the territory, since it gets rid of the size of the province and measures the degree to which the length of road kilometres occupies a given land area. On the other hand, the topography of the land can be measured by elevation or by the percentage of mountainous territory.[22] The selection of the more appropriate geographical measure is not trivial. For instance, elevation does not distinguish territories that are chiefly mountainous from territories that are hill in the majority of their extension with some more lonely mountainous zones with prominent peaks. These thoughts should be taken into consideration in the relationship between Roman roads and geography. On these bases, Table 3.1 displays the correlations between the different Roman roads measures (simple index in kilometres and in density) and two main geographical features of the territory (elevation and percentage of mountainous territory). Four diverse typologies of Roman roads are considered and the selected significance level is 5%.

Table 3.1 Correlations between Roman roads and geographical features of the territory

Roman road measures		Elevation	% Mountain
km	All	0.105*	-0.069*
	Major	0.180*	0.023
	Certain	0.108*	-0.003
	Certain & major	0.119*	0.026
km per 100 sq. km	All	-0.156*	-0.272*
	Major	-0.205*	-0.254*
	Certain	-0.119*	-0.154*
	Certain & major	-0.149*	-0.149*

Source: Author's elaborations from Istat data and McCormick, M. et al. 2013. "Roman Road Network (version 2008)", DARMC Scholarly Data Series 2013-5

The observation of Table 3.1 immediately shows how the correlation between the simple measure of Roman roads in kilometres and the elevation index is positive and significant, suggesting that more kilometres of Roman roads are needed to reach more demanding territories. When looking at measure in density, instead, correlations report negative rather than positive signs, implying that, when controlling for the concentration of roads in a defined

[22] Data on the percentage of mountainous territory are provided by Istat and are better explained in Subparagraph 3.4.2.

land area, Roman roads are sparser. These results are confirmed using the percentage of mountainous territory. Whereas for the measure in density negative and significant correlations result for all typology of Roman roads, this is not the case for the simple measure in kilometres: only the index of all Roman roads reports a significant association with the mountainous territory, and the value of the correlation is low. Findings coming from this correlation analysis are in line with the evidence provided by the literature. Roman roads, as modern roads, are less concentrated in those territory that requires more costs and energy to construct and maintain transport infrastructure. However, the high and positive correlations between the simple measure in kilometres and the elevation confirm how the Romans did not exclude demanding territories along their network.

The analysis of the literature and the results coming from both graphical and statistical analyses performed lead to identical conclusions. The Roman road network covered in a homogenous manner the entire Italian peninsula, extending in plain and mountainous areas as well. More challenging territories were not avoided by the Romans. Roads were sparser in harder terrains, but the current infrastructure, although the existence of superior technological instruments, is less concentrated in these areas as well. A more accurate and deeper investigation on the relationship between Roman roads and geographical features would need more advanced technical tools and topographical devices. However, for the purpose of this chapter, the simple analysis performed has allowed to highlight how the Romans constructed their infrastructure in a thoughtful manner, circumventing geographical obstacles when it was not possible to deal with them, but exploiting their technical knowledge to reach harder territories.

3.4 Data and descriptives

As remarked above, this chapter has a double pioneering spirit. On the one hand, it represents one of the first empirical applications that aims to explore what is the connection between the existence of the old Roman road network and current economic outcomes, exploiting on this purpose a new measure of Roman roads ad hoc constructed for this investigation. On the other hand, it is the first one, to the author's knowledge, that employs the measure of trade costs not a the country level, but at the sub-country level, looking at the pair Italian province i and world country j and proving in this way an estimate of trade costs for the Italian provinces in their trade relationship with the rest of the world. Therefore, before dealing with the empirical model, a complete descriptive analysis based on figures and statistics about provincial trade costs is required, to understand what data attitudes are and to appreciate patterns and potential phenomena pictures. In this paragraph a descriptive of the main independent variable, the measure of Roman roads in kilometres and in density, is not provided, since the extensively analysis performed in Chapter 2. Nevertheless, the paragraph gives account of all other controls that are used in the inference analysis and that can be distinguished in four groups (historical, geographical, infrastructural and productivity variables). On this aim, this paragraph has been divided into two parts: all descriptives on Italian provincial trade costs (the dependent variable of the econometric analysis) and all descriptives on all other variables (the other independent variables).

3.4.1 Italian provincial trade costs

To measure trade costs for the Italian provinces, the 'top-down' approach proposed by Novy (2013), and further investigated in Jacks et al. (2008), Chen and Novy (2011) and Chen and Novy (2012),[23] has been adopted.

Following the 'top-down' approach, the measure, based on the insights of the gravity model,[24] is aimed to capture all possible bilateral costs associated with international trade. The result is a comprehensive aggregate measure of bilateral trade costs in the form of the geometric average of international trade costs between countries i and j, relative to domestic trade costs within each country. Accordingly, the measure at the Italian provincial level has been calculated as yearly average of the geometric average of the bilateral trade costs between province i and country j, obtaining a single measure for each province and each year. Since the novelty of the measure, both descriptive and inference analysis are aimed to provide the best understanding of Italian provincial trade costs, hence trade costs are computed for a 8 years period (2003-2010).[25] Data are available for 107 out of 110 provinces. The provinces of Barletta-Andria-Trani (in south-eastern Italy, Apulia), Fermo (in east-central Italy, Marche) and Monza e della Brianza (in north-western Italy, Lombardy) have been established in 2004, but they became operative only in 2009; therefore, they are not included in the data. The same applies for the four 'new' Sardinia's provinces[26] established in 2001, but in work from 2005:

[23] Although these works report a prior date, the work of Novy of 2013 precedes them, since Dennis Novy began working on his paper some years before.

[24] Anderson and van Wincoop (2003), Anderson and Yotov (2010, 2012), Fally (2014).

[25] As it will be remarked in the empirical analysis, the use of a panel data set, rather than a cross-section one, allows to exploit more information, a higher variability and a lower collinearity among observations.

[26] Carbonia-Iglesias, Medio-Campidano, Ogliastra, Olbia-Tempio.

data are available from 2007 to 2010. Moreover, for the historical four Sardinian provinces,[27] data for 2006 are missing.

Four elements are needed to produce an estimate of trade costs according to the 'top-down' approach: i) exports from province i to country j; ii) exports from country j to province i (imports to province i from country j); iii) intra-provincial trade of province i, calculated as GDP minus province i's total exports; iv) intra-national trade of country j, computed as GDP minus country j's total exports:

$$\tau_{ij} = \left(\frac{x_{ii}x_{jj}}{x_{ij}x_{ji}}\right)^{\frac{1}{2(\sigma-1)}} - 1 = \left(\frac{(GDP_i - Tot\_EXP_i)(GDP_j - Tot\_EXP_j)}{EXP_{ij}EXP_{ji}}\right)^{\frac{1}{2(\sigma-1)}} - 1$$

Data on imports and exports at the provincial level are from the Istituto Nazionale di Statistica (Istat).[28] They refer to 192 world countries from 2003 to 2010. Since original data are expressed in euro, the conversion in US dollars has been performed using the ECB reference exchange rate US$/€ from the European Central Bank (ECB).

To calculate the internal trade by province, both GDP and total exports at the provincial level are needed. Total exports have been computed adding up all exports to single exporting country partners and converting US dollars to euro. To obtain the provincial GDP, the Italian GDP from World Bank WDI[29] database and the provincial value added (VA) from Istat have been used. The GDP from World Bank has been decomposed using the weights derived from the VA. In other terms, for each province the information on the provincial VA has been exploited to compute the share that each Italian province has on the total Italian VA, and these shares have been adopted to disaggregate the total Italian GDP in provincial GDP. Finally, the provincial internal trade has been obtained subtracting the total value of exports from the provincial GDP.

Intra-national trade has been computed using GDP reported in the World Bank WDI database and subtracting the total value of exports obtained adding up bilateral trade data from the BACI-CEPII data set revision of the Comtrade UN database.[30]

[27] Cagliari, Oristano, Nuoro, Sassari.

[28] Italian National Institute of Statistics (Istat) is the main producer of official statistics in the service of citizens and policy-makers in Italy.

[29] WDI is the acronym for World Development Indicators.

[30] CEPII (Centre d'Études Prospectives et d'Informations Internationales) develops the BACI (Base pour l'Analyse du Commerce International) data set revising the information from the Comtrade UN database.

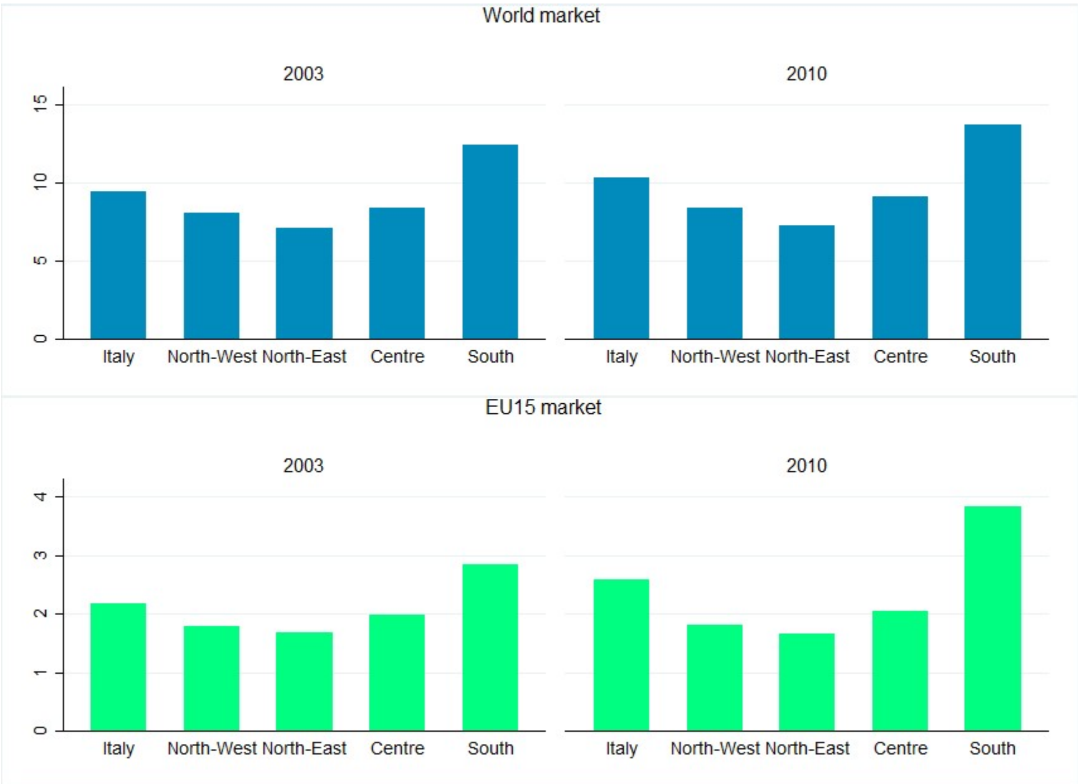
To preserve the information inferred when trade is absent, the value of zero trade (no trade) have been replaced with one; in this way it is possible to maintain observations that are important for the aim of the analysis.
The elasticity of substitution (σ), exploited in the computation of trade costs, follows what the main contributions in the literature suggest. As observed in Paragraph 1.6 in Chapter 1, Anderson and van Wincoop (2004) conclude that σ is likely to range from 5 to 10 and that, for goods that are differentiated and less substitutable, estimates are around 7 or 8.[31] In view of this, and since the intention to consider an all-inclusive set of goods, rather than a specific sector or typology of commodities, σ used in the measurement of trade costs is equal to 8. Although all descriptives refer to trade costs computed using an elasticity of substitution of 8, the robustness of estimates in the inference analysis has been checked adopting a lower (σ=7) and a higher (σ=11) elasticity.[32]
Before focusing on the descriptive analysis, one more issue on the dependent variable should be addressed. The complete description of trade costs at the provincial level, with the detail on sources and data tricks used, helps to clearer discuss what is ‘good’ and what is ‘less good’ about the measure. Until now, the empirical investigations based on the indirect measure of trade costs have concentrated on national-bilateral analyses adopting the original ‘top-down’ approach. The measure of trade costs à la Novy, at the sub-country level, is a novelty in the trade costs literature. Moreover, the focus on the only Italian province side represents a further originality in the use of the measure. If, one the hand, this chapter is a pioneer in the exploitation of the trade costs measure at the Italian provincial level, on the other hand, as highlighted in Paragraph 1.5, the measure is not free from limitations. The main constraint lies in the inaccuracy in the computation of the provincial internal trade, since it includes the ‘exports’ that each province performs outside the provincial boundaries but inside the home country. Unfortunately, the information on the ‘exports’ that each province carries out outside its borders is not available and, at the moment, a correction of the measure is not possible. A second limitation lies in the computation of the provincial GDP. Individuals' consumes and investments typically do not concern what is produced in or belongs to the province: investments are carried out outside the provincial boundaries and consumers purchase extra-provincial goods as well. Missing data on similar information makes this adjustment not feasible. Although these weaknesses cannot be addressed and corrected in this chapter, the potential of the measure à la Novy for the Italian provinces is the information it gives and

[31] Accordingly, Duval and Utoktham (2011a, 2011b) exploit an elasticity of substitution of 8.
[32] Eaton and Kortum (2002), Jacks et al. (2010) and Anderson and Yotov (2012) adopt σ=11.

what it is able to explain: the level of all tangible and intangible obstacles that allows to distinguish when for the province it is profitable to trade inside its boundaries and when it is, instead, more convenient to trade internationally.

After having examined all the key points linked to the dependent variable, a complete descriptive investigation on the indirect measure of trade costs at the provincial level is performed. The descriptive analysis has a triple nature. Firstly, differences between the world and the European (EU15) market are investigated in order to control for distance and advantages to trade with EU countries. Secondly, the long-lasting duality between northern and southern Italy is examined to consider how large these dissimilarities are. Lastly, a comparison of trade costs between the initial and the final year, for which data are available, is explored to examine the change in trade costs after the crisis. Figure 3.5 enables to assess all these three dimensions: trade market, Italian duality and situation before and after crisis.

Figure 3.5 Trade costs by exporting market, initial and final year and NUTS1 socio-economic Italian regions



Source: Author's elaborations from Istat data, BACI-CEPII data and World Bank WDI database

It is immediately clear how trade costs for the world market are considerably higher than those of the EU15 market, suggesting how lower distance, being part of the same trade union and having the same currency play a significant role in making trade less costly. When Italian provinces trade with world countries, trade costs are around 10, in the EU15 market this value is approximately 2. The differences between the starting year (2003) and the final year (2010) can be better appreciated looking at Table 3.2. Although the decrease of trade costs in last thirty years, Table 3.2 shows how the crisis produced a raise of trade costs: values in 2010 are higher than those of 2003. However, the increase is quite modest, reflecting that the boom is more likely coming from a temporary changing of the economic situation rather than from a deep transformation in the structural components of trade costs. Lastly, Figure 3.5 and Table 3.2 allow to appreciate differences across Italy. On this aim, Italy has been distinguished according to the NUTS1 disaggregation, but aggregating the southern regions with the insular ones.[33] The decomposition in NUTS1 socio-economic regions is the disaggregation that more than others allows to capture not only the Italian duality, but also the differences that emerge in different geographical parts of Italy. From Figure 3.5 and Table 3.2 clearly comes into view how northern Italy has the lowest trade costs, provinces in the Centre are halfway between a developed North and a less-developed South, and the backwardness of the southern provinces reflects in higher trade costs. Within northern Italy, North-East performs better than North-West, where trade costs are slightly higher.

Table 3.2 Trade costs by exporting market, initial and final year and NUTS1 socio-economic Italian regions

Major socio-economic Italian regions (NUTS1)	World market		EU15 market	
	2003	2010	2003	2010
Italy	9.48	10.34	2.18	2.58
North-West	8.09	8.43	1.79	1.81
North-East	7.16	7.31	1.68	1.67
Centre	8.40	9.16	1.99	2.04
South	12.45	13.77	2.85	3.84

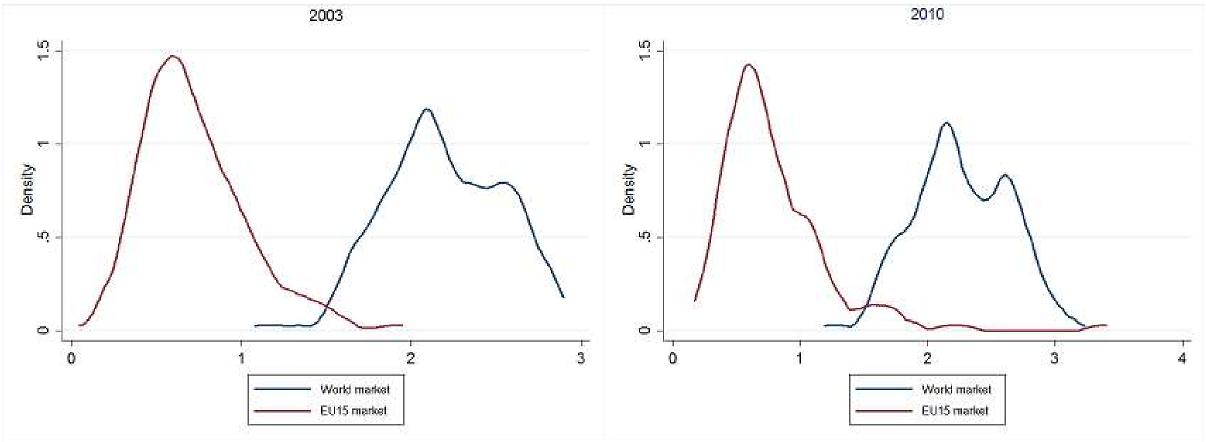
Source: Author's elaborations from Istat data, BACI-CEPII data and World Bank WDI database

To deeply explore the differences between trade costs when considering the world market and when considering the EU15 market, an investigation on the probability density functions of trade costs in both markets might be useful. On this aim, Figure 3.6 shows the Kernel distribution of the log transformation of the Italian provincial trade costs for both 2003 and 2010, using a smoothing parameter of 0.1. The bandwidth has been set in order to obtain an

[33] The North-West NUTS1 aggregation includes the following Italian NUTS2 regions: Piedmont, Aosta Valley, Liguria and Lombardy. The North-East NUTS1 aggregation comprises: Trentino-South Tyrol, Veneto, Friuli-Venezia Giulia, Emilia Romagna. In the Centre NUTS1 level are included: Tuscany, Umbria, Marche, Lazio. NUTS2 regions belonging to the South NUTS1 aggregation are: Abruzzo, Molise, Campania, Apulia, Basilicata, Calabria. In the Insular NUTS1 level are included the two main Italian islands: Sicily and Sardinia.

optimally smoothed curve close to true density. The blue line, representing trade costs for the world market, shows, in both first and last year of the data set, clear twin peaks, confirming a duality in Italian trade costs: Italian provincial trade costs for the global market are concentrated in two main ranges. When looking at trade costs referring at the EU15 market (red line), a sort of weak twin peaks appears in 2010. The right-hand side tail of the distribution is longer and thicker and presents some low peaks, suggesting that, for the EU15 market, the most of Italian provinces show uniformity in trade costs, but there is a small group of provinces for which trade costs are higher. This kind of duality feebly emerges also in 2003.

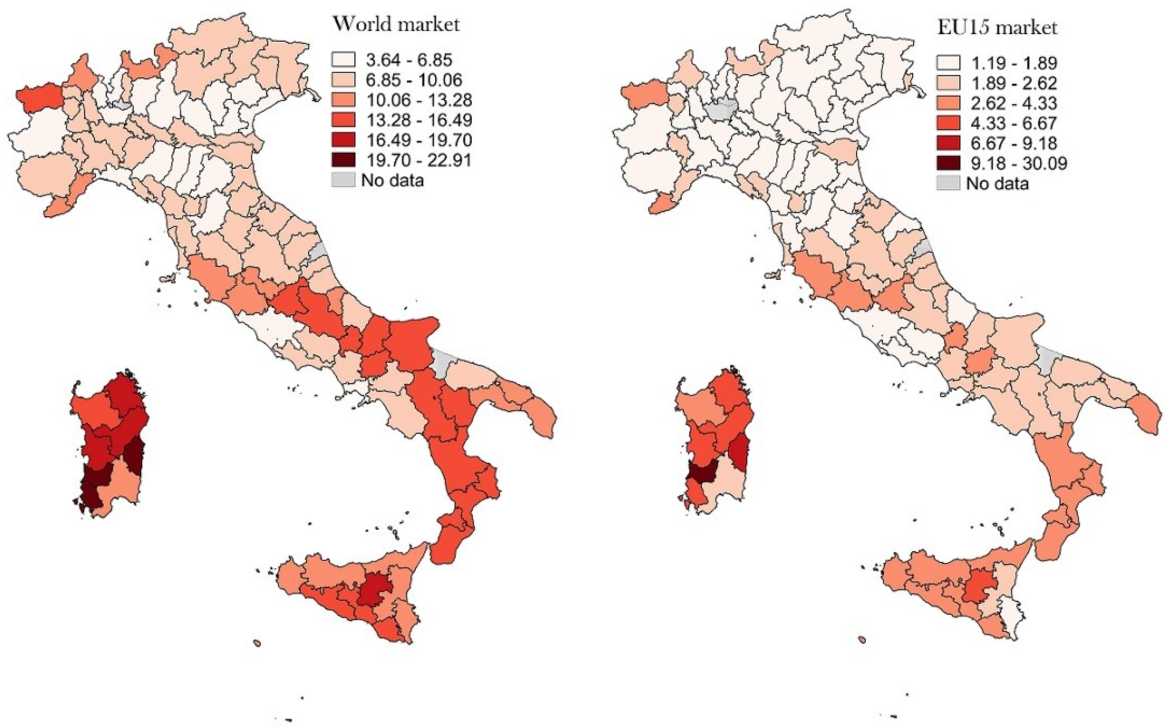
Figure 3.6 Kernel distribution of aggregated Italian provincial trade costs (in logarithms): world market and EU15 market in 2003 and 2010, bandwidth h = 0.1



Source: Author's elaborations from Istat data, BACI-CEPII data and World Bank WDI database

The duality in Italy can be better assessed with the distribution of trade costs provided in Figure 3.7. Once again, calculations for the global market and for the EU15 market have been separated to verify whether there is an effect of distance and of being included in the same trade union. Figure 3.7 reveals that trade costs in the world market present a space distribution that is connected to the level of wealth. There is a discrepancy between the northern and the southern provinces, but trade costs are also higher in provinces near the Italian border, where connectivity is still influenced by geography (i.e. the area's mountainous terrain). Moreover, Figure 3.7 highlights that this clear divergence between North and South is accentuated when considering the world, rather than the EU15 market. It is interesting to observe, how trade costs for the world market seem to be lower than those for the EU15 market. Trade costs for the world market range from a minimum of 3.64 to a maximum of 22.91; for the EU15 market they are 1.19 and 30.09. But it should be considered that just one province in Italy has, for the EU15 market, trade costs equal to 30.09; for the remaining 106 provinces, trade costs are only up to 9.17. For the world market, instead, more than half of the provinces (58 out of 107) present trade costs higher than 9.17, and more than one fourth have trade costs higher than 13, revealing that trade costs for the world market are higher than for the EU15 market.

Figure 3.7 Trade costs by Italian province in 2010, world and EU15 market



Source: Author's elaborations from Istat data, BACI-CEPII data and World Bank WDI database

This brief and preliminary descriptive on Italian provincial trade costs has produced interesting insights on the behaviour of trade costs and, more precisely, on the pattern of Italian trade costs. First of all, the analysis has confirmed, accordingly with the literature on trade costs, how distance, sharing the same currency, being part of the same economic community heavily matter in the amount of trade costs, and how these factors should be considerably taken into account in the real purpose of reducing trade costs. Moreover, the long-lasting Italian duality, extensively remarked in literature and strongly affecting numerous spheres of the economic Italian system, persists also when looking at provincial trade costs, suggesting how southern provinces need to work hard to catch the northern ones.

3.4.2 Other controls

The main independent variable of the entire chapter is represented by the measures of Roman roads both in kilometres or in density. Since the extensive analysis provided in Chapter 2, this paragraph focuses on the descriptives of the other independent variables included in the study. The other factors that potentially may affect trade costs and that typically allow to distinguish Italian provinces from each other, can be grouped into four areas: history, geography, current infrastructure and productivity.

Historical factors refer to the dominations that ruled in Italy between the twelfth and the eighteenth century and come from the work of Di Liberto and Sideri (2015). Indeed, Di Liberto and Sideri (2015) follow two approaches to measure past dominations. On the one hand, they use a set of binary variables that identify, for each province, the administration that occurred from the middle of the sixteenth century to the middle of the seventeenth century, namely, the period from 1560 to 1659. In that period, five different governments and an independent area typify the Italian peninsula, generating thus six binary variables: the Spanish Kingdom, the Republic of Venice, the Duchy of Savoy, the Papal State, the Austrians, the independent area. On the other hand, they measure the different administrations that governed Italy over seven centuries before the creation of the unified Italian State (1861), namely, the period from about 1100 to 1700, assigning to each province the number of years during which each regime has ruled. During these 700 years, nine dominations occurred: the Normans, the Swabians, the Anjou, the Spanish, the Bourbons, the Papal State, the Savoy, the Austrians and the Republic of Venice. The inference part of this chapter exploits this second set of numerical variables.

Di Liberto and Sideri (2015) find that, if a province was under the Papal State domination, the Spanish rule or the Norman government, the resulting impact on institutional quality was negative; the evidence emerging from the other dominations is more imprecise. According with what historians portray, it is possible to expect negative or positive effects from each domination. As remarked by Di Liberto and Sideri (2015), historians usually depict Normans, which ruled in the southern areas and the independent towns in the North, as having negatively affected social capital levels and, through that, development. The Swabian is identified as a positive domination: they controlled the Kingdom of Sicily (including the whole Mezzogiorno[34]) until 1266. From 1266 the Anjou conquered the southern Italy, but their administration was judged negatively, since the strong fiscal system, the regular fights

[34] 'Mezzogiorno' is an Italian term that refers to South Italy.

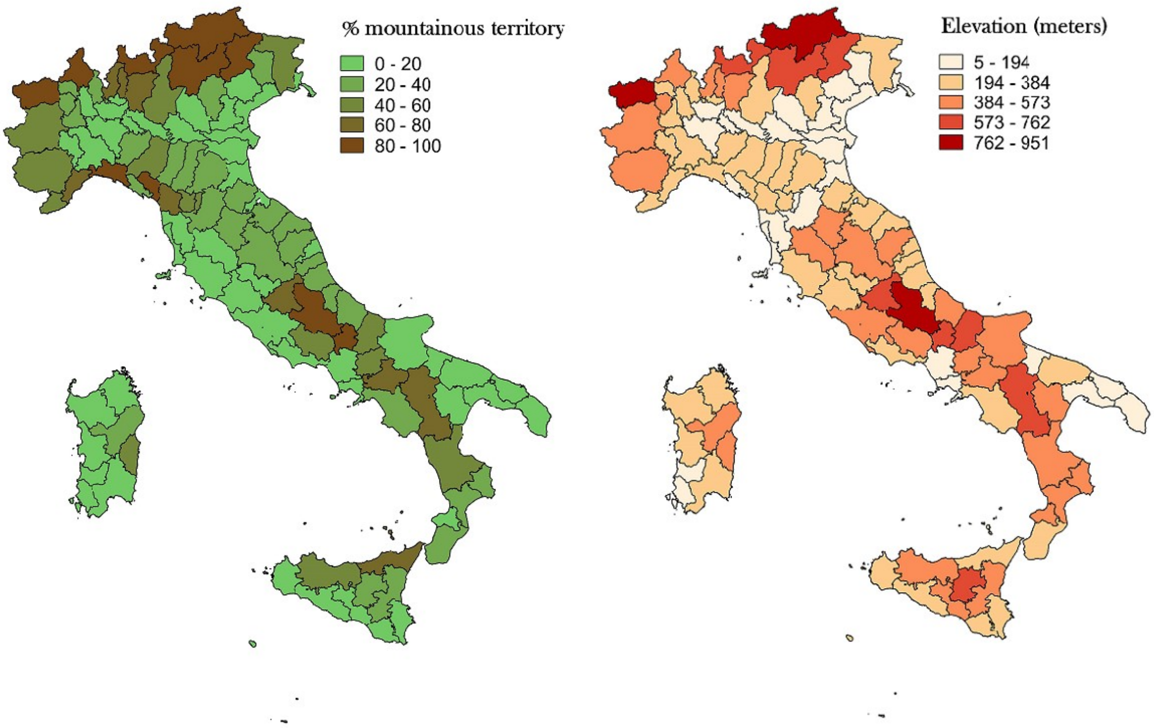
against local feudal nobility, the strict military control, the continuous wars and the abolishment of the modern state constructed by Swabians during the previous century. The Spanish kingdom influenced Italy for a long time, but, because of its inefficient institutions, bureaucracy and the implementation of extractive policies in foreign territories, the Spanish domination is depicted negatively by historians. The successors of the Spanish domination in South Italy were the Bourbons. Since the inability to improve administration in the territories inherited by the Spanish, the Bourbon domination cannot been judged as having had a positive effect on development. In the Centre of Italy, the Papal State ruled for the most part of the period examined. Although it had given evidence of good administration, it was limited to the city of Rome; in the other territories there was a diarchy between the religious and the local power. For these reasons, the Papal state is expected to have had a negative influence on institutions. The Savoy governed in the Aosta Valley, in Piedmont and in Sardinia. The effect on institutions of the Savoy domination is ambiguous: although the government was characterised by a strong central power and an authoritarian bureaucracy, the Savoy constituted a modern organisation, with the progressive passage from a feudal state to a modern one, but only in the northern territories. The Austrian domination, which dominated Italy since 1713, is considered as having positively affected institutions. They ruled first in North-East of Italy, then they conquered the Duchy of Milan, Sardinia (until 1720), the Kingdom of Naples and Sicily until 1734. They influenced also Tuscany and the Duchy of Parma and Piacenza. The Republic of Venice has been the only state to defend its independence. Because of the political stability and the economic prosperity, it should have had a positive impact on the institutional organisation (Di Liberto and Sideri, 2015).

The second set of variables included in the inferential part of this chapter refers to geographical factors. The relationship between Roman roads and geography has been investigated in deep in Paragraph 3.3. The topography of the land is fundamental in better assessing the effect of the infrastructure, since the strong correlation between landscape and transportation. In the inference part, two main geographical variables have been exploited in instrumental terms: the percentage of mountainous territory and elevation. Data on the percentage of mountainous territory at the provincial level are provided by Istituto Tagliacarne.[35] Indeed, Istituto Tagliacarne provides three statistics: percentage of mountainous territory, percentage of hill territory and percentage of flat territory. The sum of these three percentages is 100, the total provincial land. Data on elevation comes from Istat,

[35] Istituto Guglielmo Tagliacarne is an Italian foundation promoting economic culture in Italy by engaging in research and economic-statistical analysis.

which measures the elevation by municipality. The information by province has been obtained computing the simple average of all municipalities' elevations. Figure 3.8 maps both percentage of mountainous territory and elevation by Italian province. In 60 out of 110 provinces the mountainous land is over 20 percent. This result reflects in the elevation. In 55 provinces the average elevation is higher than 300 metres; in 9 provinces the elevation exceeds 600 metres. Istat categorises the Italian territory in three elevation zones (mountain, hill, level ground) and defines an area as mountainous if the territory has an elevation of 600 metres in North Italy and 800 metres in Centre-South Italy.

Figure 3.8 Geography by Italian province: percentage of mountainous territory and elevation in meters



Source: Author's elaborations from Istituto Tagliacarne and Istat data

The control for the current infrastructure is performed looking at two types of ways of transport: railroads and roads. Within current roads, both motorways and all roads are exploited. Data on kilometres of railroads by province are from Istat and refers to 2005. The information is provided for 103 out of 110 provinces, since the missing provinces have been established or became effectively operative after 2005.[36] Data on the current road network comes from Automobile Club d'Italia (ACI) and are updated to 2011. Until 2011 there was a

[36] More recent data including all 110 provinces are not available.

lack of data regarding the provision of road infrastructure in the different and comprehensive territorial levels. ACI filled up the need of more detailed data, collecting information from different sources. Data on motorways comes from AISCAT[37] and from ANAS[38]. ANAS provided also data on national interest roads. The regional roads have been identified first according to the Decree of the President of the Council of Ministers (DPCM) 21/02/2000, in accordance with the Legislative Decree (LD) n. 112 of 1998, singling out the roads not included in the highway and in the national road network and, then, thanks to the collaboration of ANAS and the Regions and Provinces or through published material. For provincial roads ACI used first the published catalogues, when available; in the remaining cases, it was conducted a survey among the provinces themselves. Only in a few cases it has not been able to find the necessary information. ACI classifies by province five different typologies of roads: motorways, regional roads, provincial roads, roads of national interest, roads 'to be classified'. The sum of all these roads gives the total extension of roads by province. Not all types of roads are included in each province (there are provinces without motorways, for example). Table 3.3 lists the length in kilometres of railroads, motorways and total roads for all Italian provinces and ranks them by total roads. The decision to look at the measures in kilometres, rather than in density, helps to better appreciate also the topography of the territory: extension of the province and its layout. The comparison between this table and Table 2.5 in Chapter 2 on the length of Roman roads by province helps to appreciate if there is a parallelism between the old infrastructure and the current one.

The fourth and last set of controls refers to productivity. Productivity has been computed dividing the total value added (the sum of agriculture, manufacture and services sectors) by the total number of workers. Both data are from Istat e refer to the eight years between 2003 and 2010.[39] Figure 3.9 maps the total productivity in Italian provinces both in 2003 and 2010, starting and final years. The map on the left-hand side shows productivity in 2003, and immediately emerges the historical duality of Italy: in the North-Centre productivity is significantly higher than in the South one. The lowest values are all concentrated in one region, Sardinia, where (in its four original provinces) the productivity is the lowest. When

[37] AISCAT is the Italian acronym for Associazione Italiana Società Concessionarie Autostrade e Trafori (Italian Association Motorways and Tunnels Dealers). It manages all those problems concerning the planning, design, construction, operation, maintenance and management of motorways and tunnels.

[38] ANAS is the Italian acronym for Azienda Nazionale Autonoma delle Strade (National Autonomous Society of Roads). It is the managing authority of the road and motorway networks in Italy.

[39] Since the unavailability of the information about the number of workers in 2003, the data on 2004 has been used also for 2003 checking whether significant variations in the number of workers in following years occurred.

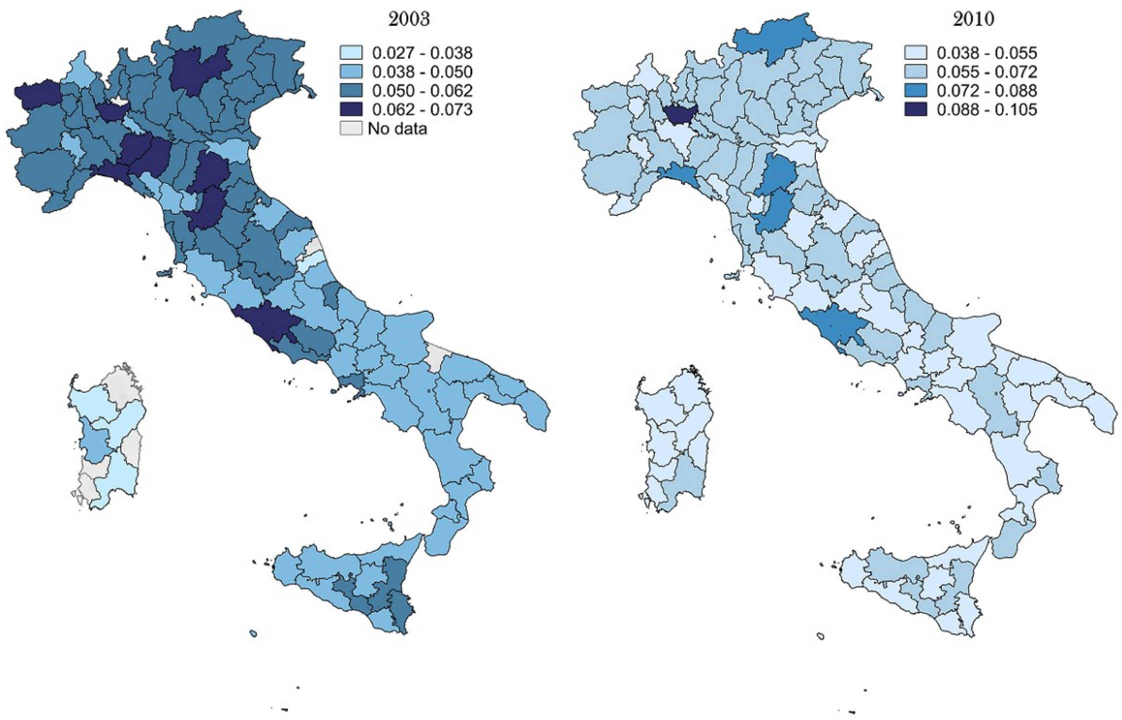
Table 3.3 Kilometres of current infrastructure by Italian province: all roads, motorways and railroads

Italian Provinces	All roads (km)	Motorways (km)	Railroads (km)	Italian Provinces	All roads (km)	Motorways (km)	Railroads (km)
Cuneo	3683	119	335	Trapani	1311	124	178
Foggia	3569	170	241	Piacenza	1309	92	97
Torino	3224	301	418	Reggio-Emilia	1293	40	37
Salerno	3202	193	274	Modena	1260	51	84
Perugia	3180	48	217	Venezia	1242	107	160
Roma	2870	332	605	Oristano	1203		82
Palermo	2624	172	231	Mantova	1180	38	107
Trento	2465	70	138	Ancona	1170	56	178
Lecce	2432		18	Olbia-Tempio	1150		
L'Aquila	2385	131	310	Belluno	1147	16	182
Alessandria	2310	181	370	Pisa	1147	42	118
Bolzano	2173	116	223	Potenza	1137	65	246
Frosinone	2162	84	227	Ferrara	1107	77	60
Pavia	2158	95	272	Isernia	1088		95
Udine	2135	151	246	Vercelli	1083	101	160
Messina	2102	197	214	Terni	1081	46	151
Chieti	2097	89	83	Milano	1069	165	368
Sassari	2015		206	Brindisi	1063		127
Grosseto	1986		125	Ascoli-Piceno	1041	44	74
Verona	1978	137	228	Ravenna	1028	48	222
Teramo	1975	89	75	Vibo-Valentia	1010	36	113
Catania	1911	95	163	Savona	998	105	143
Avellino	1906	110	175	Imperia	985	61	68
Bari	1893	78	336	Novara	957	103	399
Brescia	1852	130	163	Pescara	955	58	44
Siena	1837	61	249	Crotone	945		81
Pesaro-Urbino	1833	43	40	Napoli	944	119	157
Caserta	1816	71	295	Pordenone	943	32	93
Matera	1764		116	Fermo	912	28	
Reggio-Calabria	1679	78	190	Cremona	895	18	140
Viterbo	1679	29	198	Varese	894	46	128
Campobasso	1643	36	175	Ragusa	868		120
Firenze	1637	129	309	Massa-Carrara	825	57	76
Macerata	1629	19	93	Rovigo	777	25	106
Treviso	1608	100	217	La Spezia	773	64	81
Catanzaro	1591	47	151	Cosenza	766	138	308
Rieti	1576	29	101	Aosta	762	109	81
Siracusa	1554	58	125	Lucca	761	67	137
Parma	1552	94	179	Barletta-Andria-Trani	759	44	
Nuoro	1541		22	Verbano-Cusio-Ossola	741	18	0
Taranto	1533	23	101	Biella	715		110
Caltanissetta	1527	14	145	Livorno	715	34	0
Cagliari	1490		118	Como	668	23	146
Benevento	1482	11	141	Sondrio	606		93
Bologna	1475	172	296	Pistoia	571	29	54
Vicenza	1449	72	144	Rimini	566	30	41
Agrigento	1447		137	Lecco	541		0
Arezzo	1441	70	155	Lodi	541	39	64
Padova	1421	74	180	Medio-Campidano	488		
Bergamo	1394	32	101	Carbonia-Iglesias	475		
Asti	1376	45	185	Ogliastra	442		
Enna	1373	66	68	Gorizia	276	38	55
Latina	1351		117	Trieste	229	30	71
Genova	1322	147	210	Prato	87	10	40
Forlì-Cesena	1315	43	44	Monza e della Brianza	46	24	

Source: Author's elaborations from Istat data and from ACI (Automobile Club Italia) data, Dotazione di infrastrutture stradali sul territorio italiano, 2011

looking at 2010, this strong divergence disappears and a more uniform picture emerges. The north-central provinces are more productive than the southern ones, but the difference is not so severe. Moreover, values in 2010 are higher than those of 2003, suggesting that productivity is increased. The raise in productivity might originate from an increase of the value added, from a decrease of workers or from both (an increase of the value added and a parallel decrease of workers). But, it should be considered that the value added is at current and not at constant prices,[40] hence the potential increase of the value added might derive from a higher industry's contribution to GDP or it is simply the consequence of inflation. Looking at data, it emerges that the number of workers is decreased from 2003 to 2010. At the same time, the value added is increased. Not considering the provinces for which the information in 2003 is absent, it comes out that in Italy productivity from 2003 to 2010 is increased by 13.2 percent, the total value added is increased by 14.6 percent and the number of workers is decreased by 2.4 percent. Since it is not possible to assess, from this simple analysis, what is the origin of the augmented value added, it can be only concluded that both increased value added and decreased number of workers explain the improved productivity.

Figure 3.9 Total productivity by Italian province: 2003 and 2010



Source: Author's elaborations from Istat data

[40] At the provincial level the value added is available only at current prices.

3.5 The empirical model

The empirical approach used to explore the issue of the persistence of the Roman road network on current trade costs has followed two identification strategies. On the one hand, the first and basic interest has been to assess which is the pure and comprehensive effect of Roman roads on current trade costs, disregarding in this phase through which channel Roman roads perform and just wondering whether and what has been the impact on the modern economy. On the other hand, the attention has been devoted to the channel through which Roman roads perform, investigating whether the Roman road network left a mark on the current infrastructure and exploring, using the old Roman network as an instrument, the extent of the effect of present roads on trade costs.

The dependent variable exploited in the entire empirical analysis is represented by the (natural) logarithm of the yearly average at the Italian provincial level of the geometric average of bilateral trade costs between province i and country j. In order to avoid problems generating from zero trade flows,[41] $x_{ij}=0$ or $x_{ji}=0$ have been replaced by 1. This approach is the one proposed by McCallum (1995) and Raballand (2003), who suggest to replace the zeros with small positive numbers. In the gravity model framework the case of zero trade flows is serious, since usually the dependent variable is expressed in logarithms and the logarithm of zero is undefined. As highlighted by Haq et al. (2011), beyond the method suggested by McCallum (1995) and Raballand (2003), researchers have adopted different strategies to deal with this problem. McCallum (1995) and Frankel (1997) suggest to delete zeros, but the omission of zeros generates sample selection bias. Rose (2000) proposes to estimate a Tobit model and censor the zeros at the left tail. Helpman et al. (2008), Emlinger et al. (2008), Disdier and Marette (2010), Jayasinghe et al. (2010) use a Heckman selection model. Literature has highly discussed the advantages and the weaknesses behind each approach. The one selected in this chapter[42] is the most common in empirical research. The final measure of indirect trade costs adopted in the inference analysis is, hence, a yearly average at the provincial level, which originates from the geometric average of bilateral measure of trade costs between each province i and each country partner j, where zero trade flows have been substituted with 1, obtaining one measure for each Italian province for each year between 2003 and 2010.

[41] According to Wooldridge (2006), if data are not randomly missing or not randomly zero (as for the gravity model) but the sample selection procedure uses an exogenous sampling rule, then OLS estimates could be unbiased. On the contrary, if the sampling rule is endogenous (as it is in the gravity model environment, where only trade flows greater than zero are considered), then OLS estimates are biased.

[42] It is the approach also followed in the robustness checks of Chapter 4 for country trade costs.

Table 3.4 Data and sources

Variables	Definition	Years	Source	Other info
Trade costs	Yearly average at the Italian provincial level of the geometric average of international trade costs between Italian province i and country j relative to domestic trade costs within each province	2003-2010	Istat, BACI-CEPII and WDI	107 provinces
All Roman roads (km)	kilometres of all Roman roads	117 A.D.	Author's creation from McCormick et al. (2013) shape file	110 provinces
Certain Roman roads (km)	kilometres of certain Roman roads	117 A.D.	Author's creation from McCormick et al. (2013) shape file	110 provinces
Major Roman roads (km)	kilometres of major Roman roads	117 A.D.	Author's creation from McCormick et al. (2013) shape file	110 provinces
Certain and major Roman roads (km)	kilometres of certain and major Roman roads	117 A.D.	Author's creation from McCormick et al. (2013) shape file	110 provinces
All Roman roads (density)	kilometres of all Roman roads per 100 square kilometres of land area	117 A.D.	Author's creation from McCormick et al. (2013) shape file and Istat	110 provinces
Certain Roman roads (density)	kilometres of certain Roman roads per 100 square kilometres of land area	117 A.D.	Author's creation from McCormick et al. (2013) shape file and Istat	110 provinces
Major Roman roads (density)	kilometres of major Roman roads per 100 square kilometres of land area	117 A.D.	Author's creation from McCormick et al. (2013) shape file and Istat	110 provinces
Certain and major Roman roads (density)	kilometres of certain and major Roman roads per 100 square kilometres of land area	117 A.D.	Author's creation from McCormick et al. (2013) shape file and Istat	110 provinces
Normans	Number of years of the Norman domination	1100-1700	Di Liberto and Sideri (2015)	110 provinces
Swabians	Number of years of the Swabian domination	1100-1700	Di Liberto and Sideri (2015)	110 provinces
Anjou	Number of years of the Anjou domination	1100-1700	Di Liberto and Sideri (2015)	110 provinces
Spain	Number of years of the Spanish domination	1100-1700	Di Liberto and Sideri (2015)	110 provinces
Bourbons	Number of years of the Bourbon domination	1100-1700	Di Liberto and Sideri (2015)	110 provinces
Papal State	Number of years of the Papal domination	1100-1700	Di Liberto and Sideri (2015)	110 provinces
Venice	Number of years of the Venetian domination	1100-1700	Di Liberto and Sideri (2015)	110 provinces
Austria	Number of years of the Austrian domination	1100-1700	Di Liberto and Sideri (2015)	110 provinces
Savoy	Number of years of the Savoy domination	1100-1700	Di Liberto and Sideri (2015)	110 provinces
Mountain	Percentage of mountainous territory	time invariant	Istituto Tagliacarne	110 provinces
Elevation	Elevation in metres	time invariant	Istat	110 provinces
Railroads (km)	Kilometres of railroads	2005	Istat	103 provinces
Motorways (km)	Kilometres of motorways	2011	Automobile Club d'Italia (ACI)	110 provinces
All current roads (km)	Kilometres of all current roads	2011	Automobile Club d'Italia (ACI)	110 provinces
Productivity	Total value added / total workers	2003-2010	Istat	110 provinces

Source: Author's elaborations

The entire empirical analysis focuses on three pillars. First, it exploits panel data in order to benefit from more variability and more precise estimates. Second, it refers to the world market in order to take advantage from more heterogeneity, i.e. Italian provincial trade costs have been computed considering trade with world countries partners. Third, the elasticity of substitution used to compute trade costs is constant and equal to 8, according to the main evidence coming from the literature. The robustness checks call into question these three pillars and test the solidity of estimates, using cross-section data rather than panel data, looking at the EU15 market rather at the world market and adopting both a lower (σ=7) and a higher (σ=11) elasticity of substitution.

The measure of Roman roads is the main independent variable of the entire analysis, but other factors are controlled for in order to obtain consistent estimates. These factors can be grouped into four categories: history, geography, current infrastructure and productivity. Table 3.4 summarises the data and the sources used to construct all the variables included in the analysis.

3.5.1 The direct effect

The identification strategy is aimed at assessing the simple and broad *direct effect* of the Roman road network on current trade costs, ignoring what is the mechanism that enabled the Roman roads to produce persisting effects that lasted until today, but just focusing on the widespread legacy of the old Roman infrastructure. On this aim, the inference analysis started from an original plain and compact empirical model to further extend it in order to control for all those potential factors that might affect current trade costs. In other terms, the identification strategy has proceeded by stages, complicating stage by stage the model.

The first step and, therefore, the first empirical model is what can be called the ***base model***. It consists in estimating a compact model where the dependent variable is represented by the natural logarithm of the indirect measure of trade costs at the Italian provincial level, the main independent variable is the natural logarithm of the Roman road measure in kilometres and the other covariates are represented by a series of historical variables. Regional fixed effects are added in order to avoid collinearity problems:

$$\ln \tau_{it} = \beta_0 + \beta_1 \ln RR_i + \beta_2 \boldsymbol{H}_i + \eta_s + \gamma_t + \varepsilon_{it}$$

where:

- τ_{it} denotes the dependent variable, the yearly average of all trade costs at the Italian provincial level from 2003 to 2010
- RR_i denotes the measure of Roman roads by province
- $\mathbf{H}_i$ denotes the vector of historical measures (past dominations) by province
- η_s denotes regional fixed effects
- γ_t denotes time effects
- ε_{it} denotes the error term

Different typologies of Roman roads have been adopted: all, certain, major, both certain and major. The historical variables refer to the nine dominations that occurred in Italy after the fall of the Roman empire, until the unification of Italy. They are expressed in number of years and help to control for the historical administrations that ruled in Italy from the Middle Ages to the modern times. Since the time-invariant nature of the Roman road measure, it is not possible to perform a fixed effects model at the provincial level, hence regional fixed effects are added to control for the variability at least at the NUTS2 level. Time effects and the error term complete the analysis.

The assumption underlying the fixed effects model is that the individual effect for each province i is fixed and that the variation is within the panel variable (the province i). The use of the within estimator allows to assess this variation, since it estimates the model, exploiting for each regressor the deviation from the individual average value. In the fixed effects model, the individual effects exhibit in intercepts: each observation has a different intercept. The estimates obtained with this model are consistent and unbiased even if the independent variables are correlated with the error term. Unfortunately, since the within transformation drops out all those variables that do not vary over time and since the main independent variable, Roman roads, is time-invariant, it is not possible to exploit such a model. A solution is to rely on regional fixed effects.

The measure in kilometres of Roman roads considers implicitly also the topography and the size of the province: more rugged or bigger provinces should include a higher number of kilometres of Roman roads. In order to take into account these features, the analysis is deepened exploiting the measure of Roman roads also in density: kilometres of Roman roads per 100 square kilometres of the land area. In this way, it is possible to control for the size and the landscape of the land. The model that employs the Roman road measure in density has the identical structure of the base model above and can be named ***model in density***.

The third step of the 'direct effect' strategy adds a second control: the current infrastructure. According to Kessides (1993), and as underlined in Paragraph 3.2, current infrastructures affect economic development and performance in a large number of ways, acting through different channels and including externalities, spillover effects and indirect mechanisms that must be taken into account. In this framework, assuming that better economic outcomes are determined by denser current road network, and not vice versa, is a very strong assumption. Literature highly supported the idea according to which roads and railways are often built to connect already developed regions, but regions can reach development after the construction or improvement of infrastructures. It is more likely that rich countries/regions/provinces can afford better infrastructures. Since the potential endogeneity problem that might rise simply including the current infrastructure in the regression model, the strategy exploited is the one proposed by Ramcharan (2009) and Del Bo and Florio (2012). Del Bo and Florio (2012) argue that the instrumental variables (IV) approach is a suitable method to overcome the possible endogeneity problem. On this aim, following Ramcharan (2009), they use geographical variables as an instrument for the current infrastructure, relying on the percentage of arable land, of mountainous terrain and the length of natural watercourses. Ramcharan (2009) highlights how the topography of the territory is strictly correlated with the land development of infrastructure. He argues that current roadways and railways belong to economic choices, and the potential endogeneity problem might bias OLS estimates. On this aim, he suggests that geographical variables are good instruments to better explore transport infrastructure.[43] Following this literature, the two-stage least square (2SLS) approach[44] on a panel data set has been adopted relying on the percentage of mountainous territory and on elevation as instrumental variables for the current infrastructure. Both instrumental and instrumented variables have been expressed in natural logarithms. The current infrastructure (CI), that is controlled for in the model, is represented by the kilometres of railroads, kilometres of motorways and kilometres of total roads. The model is completed by the

[43] Montolio and Solé-Ollé (2009) rely also on the instrumental variable approach to avoid endogeneity issues, but they focus on political variables. Political variables are suitable instruments in specific cases. Investigating on the public investment in the road infrastructure for Spanish provinces as a determinant of the total factor productivity growth, Montolio and Solé-Ollé argue that in Spain public investment decisions are taken by the central government, not by the provinces, hence political variables satisfy the restrictions imposed by the instrumental variable approach.

[44] In Paragraph 3.5.2 a brief explanation of the instrumental variable (IV) and of the 2SLS approaches is provided.

essential measure of Roman roads in kilometres, historical controls, regional fixed effects, time effects and error term. The first and second stages are:

First stage

$$\ln CI_i = \alpha_0 + \alpha_1 \ln \boldsymbol{G}_i + \alpha_2 \ln RR_i + \alpha_3 \boldsymbol{H}_i + \eta_s + v_i$$

Second stage

$$\ln \tau_{it} = \beta_0 + \beta_1 \ln RR_i + \beta_2 \ln \widehat{CI}_i + \beta_3 \boldsymbol{H}_i + \eta_s + \gamma_t + \varepsilon_{it}$$

Since geographical variables are used as instruments for the current infrastructure, the model above can be named ***geography for infrastructure model***.

The fourth and last step adds a further control and includes in the model the total productivity (P). On this aim, the above model is extended comprising the natural logarithm of the productivity:

First stage

$$\ln CI_i = \alpha_0 + \alpha_1 \ln \boldsymbol{G}_i + \alpha_2 \ln RR_i + \alpha_3 \boldsymbol{H}_i + \alpha_4 \ln P_i + \eta_s + v_i$$

Second stage

$$\ln \tau_{it} = \beta_0 + \beta_1 \ln RR_i + \beta_2 \ln \widehat{CI}_i + \beta_3 \boldsymbol{H}_i + \beta_4 \ln P_{it} + \eta_s + \gamma_t + \varepsilon_{it}$$

It must be underlined that the control for the current productivity might be endogenous, since the potential correlation with the error term.[45] Aware of this problem, the fourth model, that can be named ***extended model***, has the precise intent to verify if the Roman road coefficient is robust and persists also with a similar control.

In sum, the identification strategy, focused on investigating the direct effect of the Roman road network, starts from a base model which simply uses the Roman road measure in kilometres and controls for past dominations. To control for the strength of the Roman road measure, Roman roads are exploited in density rather than in kilometres. In a further step, the control for the current infrastructure is added through an instrumental variables approach that relies on geographical variables (mountainous territory and elevation) as instruments for

---

[45] "*If any one regressor is endogenous then in general OLS estimates of all regression parameters are inconsistent (unless the exogenous regressor is uncorrelated with the endogenous regressor)*" (Cameron and Trivedi, 2005, p. 92).

present kilometres of railroads, motorways and total roads. Last step consists in including the productivity as a further control for the robustness of the estimates.

3.5.2 The indirect effect

The second identification strategy goes much more in deep, assessing whether the old Roman infrastructure produces a persistent effect on the current measure of trade costs via the current infrastructure. This second approach is aimed at the channel through which Roman roads perform and is more linked to the *indirect effect* of the historical infrastructure.

The strategy followed to assess the indirect effect of the Roman road network through current roads refers to the instrumental variables (IV) approach and the two-stage least square (2SLS) estimator. These methods basically consist in using and relying on an instrumental variable to produce consistent estimates. The IV approach is normally adopted to address the classic problems caused by the use of endogenous regressors and that typically arise in Ordinary Least Square (OLS) regressions: omitted variable bias, measurement error, reverse causality. The current infrastructure is typically affected by endogeneity and the use of instrumental variables methods helps to obtain consistent parameter estimates. As highlighted by Cameron and Trivedi (2005), this approach is frequently exploited in econometrics, but rarely used elsewhere, since it is conceptually difficult. If, on the one hand, the advantage of the IV methods basically is based on addressing omitted variable bias, measurement error or reverse causality problems, on the other hand, the main challenge lies in two important conditions that need to be fulfilled: the correlation between instrumental variable and instrumented variable should be different from zero, the correlation between instrumental variable and the error term should be zero. These two conditions, that are unavoidable to exploit instrumental variables methods and that define the *validity* of an instrument, represent the key to assess the effects of the Roman roads network on current trade costs through the current roads. The first and second stage of the IV approach, where the Roman road measure represents the instrumental variable and the current road variable (CR) represents the instrumented variable, can be written in the following way:

First stage

$$\ln CR_i = \alpha_0 + \alpha_1 \ln RR_i + \alpha_2 \boldsymbol{H}_i + \alpha_3 \ln P_i + \eta_s + v_i$$

Second stage

$$\ln \tau_{it} = \beta_0 + \beta_1 \ln \widehat{CR}_i + \beta_2 \boldsymbol{H}_i + \beta_3 \ln P_{it} + \eta_s + \gamma_t + \varepsilon_{it}$$

The validity of the Roman road instrument is given by the following two conditions: Cov (CR, RR) ≠ 0 in first stage; Cov (RR,ε) = 0 in second stage. The first condition refers to the correlation between the old infrastructure and the current infrastructure and is easily verifiable through the first stage statistics. The key idea behind this is that territories with a denser Roman road infrastructure are more likely to have a denser present road infrastructure. The second condition requires that the instrument is uncorrelated with any other determinant of the dependent variable. While we can test whether the first condition is satisfied, the second condition cannot be tested when the model is exactly identified (number of instruments equal to the number of endogenous regressors). Over-identified models allow instrument exogeneity to be tested. In short, the challenge of employing instrumental variables methods is finding valid instruments; an instrument is valid when it is relevant (i.e. it fulfils the first condition) and exogenous (i.e. it fulfils the second condition). It is challenging to find variables that meet the definition of valid instruments: conceptually, most variables that have an effect on endogenous variables may also have a direct effect on the dependent variable. The exogeneity of the Roman road instrument is corroborated by Garcia-López et al. (2015) and by De La Roca and Puga (2017), in two studies of Spain about urbanisation and agglomeration economies. Garcia- López and co-authors exploit the Roman road measure in kilometres as an instrument for modern highways to explore the suburbanisation of 123 metropolitan cities. De La Roca and Puga (2017) adopt the number of Roman road rays located within 25 km from each urban centre as an instrument for current city sizes.

3.6 Estimation results and robustness checks

This paragraph has a twofold scope. On the one hand, it has the aim to present the results coming from the empirical estimations and to discover whether the Roman road network affects current Italian provincial trade costs. On the other hand, it gives the opportunity to think about the functioning and the power of the Roman road measure and to exploit it in new or improved empirical applications.

The identification strategy has been divided into two parts. The identification strategy, that has been named ‘direct effect’ considers the pure and comprehensive effect of Roman roads, controlling for history, current infrastructure and productivity. The second identification wonders about the channels through which Roman roads perform, questioning whether the effect on current trade costs occurs via current roads; for this reason, it has been named ‘indirect effect’.

Table 3.5 Estimation results base model: Roman roads in kilometres and density (panel data, world market)

Dependent variable: Average trade costs (ln), σ=8	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	km				**density**			
	Certain	All	Major	Certain & major	Certain	All	Major	Certain & major
Roman roads (ln)	-0.109***	-0.109***	-0.075**	-0.110***	-0.117***	-0.085*	-0.052	-0.117***
Normans	-0.009***	-0.007***	-0.007***	0.000	-0.007***	-0.004**	-0.004**	0.003
Swabians	0.001	0.000	0.000	0.001	0.000	0.000	-0.000	0.000
Anjou	-0.001	-0.000	-0.000	-0.001***	-0.001*	-0.001	-0.001	-0.001***
Spain	-0.000	-0.000	-0.000	-0.001*	-0.000	-0.000	-0.000	-0.001**
Bourbons	0.009***	0.010***	0.010***	0.012***	0.009***	0.009***	0.009***	0.012***
Papal State	0.000	0.000	0.001*	0.001	0.000	0.000	0.001*	0.001*
Venice	-0.000	-0.000	-0.001	-0.001	-0.001**	-0.001*	-0.001*	-0.001**
Austria	-0.000	-0.000	-0.000	0.000	0.000	0.000	0.000	0.000
Savoy	-0.000	-0.001	-0.000	-0.001***	-0.001	-0.001	-0.001	-0.002***
Constant	3.449***	3.457***	2.219***	2.308***	3.289***	3.209***	1.857***	1.771***
Regional fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Time effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	664	804	736	608	664	804	736	608
R-squared	0.749	0.719	0.680	0.748	0.741	0.697	0.665	0.738

Note: Asterisks denote significance levels; * $p<0.10$, ** $p<0.05$ and *** $p<0.01$.

Table 3.5 displays estimation results of both base model and model in density of the ‘direct effect’ strategy. They simply assess the pure effect of Roman roads, controlling for past dominations. The measure of Roman roads is expressed in kilometres in the first four specifications and in density in the last four regressions. Four different typologies of roads have been exploited: certain, all, major, both certain and major. Clustered standard errors at the provincial level have been included in all specifications.

Significance and negative sign of the coefficient associated to the Roman road measure persists in almost all the specifications, independently on how the Roman road measure has been computed, showing that lower trade costs are linked to a longer or denser road network. The effect of past dominations is very weak and for some administrations (Swabian, Austrian) it is completely absent. The highest effect among all governments is the one played by the Bourbons' and the Normans' dominations. Accordingly with historians' portraits and with Di Liberto and Sideri (2015) findings, the effect of the Bourbons on trade costs is negative (more years under the Bourbon domination, higher current trade costs). The Papal State, the Venetian and the Savoy dominations as well reflect what history depicts. The Norman, the Anjou and the Spanish dominations report an opposite sign from what historians describe. But, what is really notable from Table 3.5 is the strong effect played by the Roman road measure compared to past and more recent dominations. The coefficient of the Roman road measure is significant at 1% level in 5 out of 8 specifications; just in one regression the coefficient is not significant, and its effect on current trade costs is high. Finally, the R-squared values inform about a satisfying goodness-of-fit in both kilometres and density models. Estimates in Table 3.5 refer to the world market,[46] have been obtained using panel data, and trade costs have been computed using an elasticity of substitution equal to 8. Table 3.8 and Table 3.9 in Appendix 3.2 display estimates for the base model and the density model as well, but performing some variations. Estimates in Table 3.8 refer to the EU15 market[47] and the dependent variable of trade costs has been obtained using a higher (σ=11) elasticity of substitution. In Table 3.9, cross-sectional data have been exploited and trade costs have been computed adopting a lower sigma (σ=7). From these robustness checks immediately emerges the large negative effect of the Roman road network on current trade costs and how this effect persists disregarding the typology of data exploited, the reference market and the elasticity of substitution used to obtain trade costs.

[46] Italian provincial trade costs have been computed considering all world country partners of each province.
[47] Italian provincial trade costs have been computed considering only EU15 country partners of each province.

Table 3.6 Estimation results IV approach for current infrastructure (geographical variables as instruments) and check for productivity (panel data, world market)

Dependent variable: Average trade costs (ln), σ=8	(1)	(2)	(3)	(4)	(5)	(6)
Certain Roman roads in km (ln)	-0.072***	-0.083***	-0.085***	-0.053***	-0.054**	-0.062***
Total roads in km (ln)	-0.370***			-0.364***		
Motorways in km (ln)		-0.098*			-0.081	
Railroads in km (ln)			-0.104**			-0.102**
Total productivity (ln)				-0.367***	-0.577***	-0.443***
Normans	-0.016***	-0.030***	-0.011***	-0.015***	-0.030***	-0.009***
Swabians	0.001	0.002***	0.000	0.001	0.001***	-0.000
Anjou	0.002**	-0.000	-0.000	0.002**	-0.000	-0.000
Spain	0.000*	0.000***	-0.000*	0.000	0.000**	-0.000**
Bourbons	0.010***	0.041***	0.009***	0.009***	0.042***	0.008***
Papal State	-0.000	0.000***	0.000***	-0.000	0.000***	0.000***
Venice	0.000	0.000***	-0.001***	0.000	0.000***	-0.001***
Austria	-0.000	0.000***	0.000***	-0.000	0.000***	0.000***
Savoy	0.001*	0.000	-0.000***	0.001*	0.000	-0.000***
Constant	6.185***	3.284***	3.962***	5.002***	1.394***	2.578***
Regional fixed effects	YES	YES	YES	YES	YES	YES
Time effects	YES	YES	YES	YES	YES	YES
Observations	496	432	488	496	432	488
R-squared	0.855	0.882	0.858	0.863	0.895	0.868
First-stage instrument 1 (% mountain) coefficient	0.087***	-0.054	0.244***	0.087***	-0.058	0.244***
First-stage instrument 2 (elevation) coefficient	-0.330***	-0.547***	-0.386***	-0.327***	-0.545***	-0.381***
First-stage F-statistic	18.110	14.680	28.110	18.420	14.730	39.880
Sargan P-value	0.513	0.056	0.001	0.527	0.005	0.001
Durbin-Wu-Hausman P-value	0.367	0.063	0.095	0.261	0.100	0.227
Kleibergen-Paap Wald F-statistic	18.110	14.680	28.110	18.420	14.730	39.880
Montiel-Pflueger effective F-statistic	11.738	19.092	31.262	12.409	19.275	40.829
Pagan-Hall P-value	0.000	0.000	0.000	0.000	0.000	0.000
Instruments	mountainous territory and elevation (ln)	mountainous territory and elevation (ln)	mountainous territory and elevation (ln)	mountainous territory and elevation (ln)	mountainous territory and elevation (ln)	mountainous territory and elevation (ln)

Note: Asterisks denote significance levels; * p<0.10, ** p<0.05 and *** p<0.01.

The check for the current infrastructure in the base model has been performed relying on the IV approach and using geographical variables as instrumental variables for the present infrastructure. Table 3.6 shows estimation results coming from this analysis. All six specifications refer to panel data, and use as dependent variable Italian provincial trade costs for the world market obtained using an elasticity of substitution equal to 8.

The first three columns do not include the further control for the current productivity, the last three columns do. In all regressions the instruments are represented by the percentage of mountainous territory and by the elevation. What changes is the instrumented variable: in first and fourth specifications the current infrastructure is represented by the length in kilometres of total roads; in second and fifth columns, by the length of motorways; in third and sixth, by that of railroads. The main independent variable in all regression is represented by certain Roman roads. Historical controls are included, and clustered standard errors at the provincial level have been comprised.

Table 3.6 shows highly significant, and with the expected negative sign, coefficients for both Roman roads and current infrastructure measures, confirming how both past and current infrastructure have a positive effect in reducing trade costs and, consequently, how provinces with a more intense network should be more prone to trade internationally than domestically. When considering total roads, their effect is higher than that of certain Roman roads and this is quite comprehensible, since the different period they refer to. Nevertheless, the coefficient of the motorways variable is not significant in one specification and not highly significant in the second regression, and its effect is low. Railroads seem to perform better than motorways but worse than total roads. Due the particular approach according to which these estimates have been obtained, a simple comment on signs and significance is not sufficient. A series of post-estimation tests are needed. In fact, since the strong assumptions needed to perform 2SLS, the IV analysis requires a set of tests to control for the issues that can occur using instrumental variables. Basically, heteroskedasticity, relevance and exogeneity of instruments[48] and endogenous regressors are the four main problems that should be detected after a IV estimation. Appendix 3.1 addresses all these issues by presenting a complete and extensive report on all controls and tests that should be performed after the 2SLS estimation and that are displayed in the bottom part of Table 3.6. Here the key results are briefly discussed.

[48] Relevance condition and exclusion restriction condition typify the validity of an instrument.

The main issue linked to the use of IV approaches refers to the importance of chosen instruments (not weak instruments). First-stage F-statistics prove the validity of the relevance condition. Accordingly, first-stage estimates of Table 3.6, namely the effect of geographical variables (percentage of mountainous territory and elevation) on the current infrastructure variables, confirm the selected strong instruments. The effect of elevation on current infrastructure is always negative and highly significant, according to what the literature suggests:[49] a 1 percent increase in elevation is associated with a 1 percent decline in the number of kilometres of current infrastructure. The effect of the percentage of mountainous territory on total roads and railroads is positive rather than negative, and it is absent for motorways. The positive sign of the coefficient linked to the mountainous territory variable suggests that the topography of the land matters: although elevation affects negatively the construction of transport infrastructure, it is also true that, in mountainous territory, it should exist more kilometres of transport networks, since the landscape of the land. The missing effect in case of motorways confirms this reasoning: in mountainous territory the presence of motorways is rare or totally absent. The validity of the exclusion restriction, the second typical concern in IV analyses, is checked by the Sargan's test. Results displayed in Table 3.6 suggest that the exclusion restriction condition is perfectly satisfied when the instrumented variable is represented by the kilometres of total roads; when the endogenous regressors are motorways or railroads, the overidentifying restrictions test warns about the validity of the exclusion restrictions. These results should be taken cautiously when using panel data. A further control refers to the suitability of the IV method. The Durbin-Wu-Hausman (DWH) test concludes that the variables being instrumented are endogenous in specification 1, 2, 4 and 5 of Table 3.6, confirming that the IV approach is needed. In regressions 3 and 6 the use of the IV estimator is questioned, since the null hypothesis that the instrumented variable is exogenous is not rejected. However, the detected heteroskedasticity may impact on the consistency of test results.

The picture that emerges from Table 3.6 is that geographical variables are valid instruments for the current infrastructure. Total roads are the endogenous variable that better allows to appreciate the effect of the current infrastructure on trade costs. However, what really emerges from Table 3.6 is the effect of the past infrastructure, Roman roads. Although the control for the current infrastructure and for productivity (that shows a highly significant and

[49] Physical obstacles make it harder to build transport infrastructure (Del Bo and Florio, 2012). Countries with rougher surfaces have less dense transport infrastructures (Ramcharan, 2009).

Table 3.7 Estimation results IV approach for current infrastructure (Roman roads as instruments) and check for productivity (panel data, world market)

Dependent variable: Average trade costs (ln), σ=8	(1)	(2)	(3)	(4)
Total roads in km (ln)	-0.565***		-0.367***	
Motorways in km (ln)		-0.299***		-0.169***
Total productivity (ln)			-0.875***	-1.092***
Normans	-0.018***	0.010***	-0.013***	0.004*
Swabians	0.001**	0.001	0.000	-0.000
Anjou	0.003***	-0.001***	0.002***	-0.001***
Spain	0.000	0.000***	0.000	0.000***
Bourbons	0.015***	0.080***	0.010***	0.063***
Papal State	-0.000***	0.001***	-0.000**	0.001***
Venice	0.000**	0.000**	0.000*	0.000**
Austria	-0.001***	0.000	-0.001***	0.000***
Savoy	0.000	0.000	0.000	0.000
Constant	7.261***	-2.314***	3.227***	-4.384***
Regional fixed effects	YES	YES	YES	YES
Time effects	YES	YES	YES	YES
Observations	664	528	664	528
R-squared	0.643	0.805	0.769	0.872
First-stage instrument 1 (certain RR) coefficient	0.123***	0.221*	0.120***	0.012
First-stage instrument 2 (all RR or certain & major RR) coefficient	0.140***	0.205	0.140***	0.328***
First-stage F-statistic	66.520	107.800	48.550	51.690
Sargan P-value	0.004	0.000	0.066	0.038
Durbin-Wu-Hausman P-value	0.000	0.001	0.000	0.694
Kleibergen-Paap Wald F-statistic	66.520	107.800	48.550	51.690
Montiel-Pflueger effective F-statistic	81.557	79.476	65.132	47.907
Pagan-Hall P-value	0.000	0.000	0.000	0.000
Instruments	Certain RR km (ln) All RR km (ln)	Certain RR km (ln) Certain & major RR km (ln)	Certain RR km (ln) All RR km (ln)	Certain RR km (ln) Certain & major RR km (ln)

Note: Asterisks denote significance levels; * p<0.10, ** p<0.05 and *** p<0.01.

with the expected negative sign coefficient), the effect of Roman road measure persists in all regressions.

The second part of the inference analysis has considered what has been named the 'indirect effect' strategy, i.e. the effect of the Roman infrastructure via the current infrastructure. On this aim, the Roman road measure has been used as instrumental variable relying on the IV approach. From Table 3.7 emerges that Roman roads are relevant and not weak instruments for the current infrastructure.

The high values of the first-stage F-statistic and of the Montiel-Pflueger effective F-statistic confirm this conclusion. Unfortunately, the Sargan's test on the exclusion restriction seems to validate that Roman roads are correlated with the error term. Since the indecision about the power of the test for overidentifying restrictions with panel data, Table 3.10 in Appendix 3.2 performs the same inference analysis of Table 3.7, but exploiting cross-sectional rather than panel data. In that case, the P-value of the Sargan's test is very high and does not allow to reject the null hypothesis that the excluded instruments are valid instruments. Moreover, the detection of heteroskedasticty with panel data, leads to distrust the Sargan's test results of Table 3.7. The DWH test confirms that current roads are endogenous variables and, hence, that IV methods are needed to deal with them. The first-stage coefficients of Roman road measures show how the impact of the historical Roman road infrastructure is positive, as expected, and relevant in determining the current infrastructure. The figure emerging from Table 3.7 is that Roman roads represent powerful instruments for the current infrastructure, they have a high impact on the current infrastructure, which, in turn, affects current trade costs. These results are confirmed controlling for past dominations and current productivity. With cross-sectional data (Table 3.10), first-stage F-statistic does not pass the rule of thumb; the reduced observations warn about a suited variability of data. Nevertheless, the first specification of Table 3.10 provides a satisfactory image.

From this twofold empirical analysis, the main conclusion that clearly comes into sight is the persistent and negative effect of the historical Roman road network on Italian provincial trade costs, that resists different types of controls, different data and different estimation methods.

3.7 Concluding remarks

This chapter took into consideration a novel and appealing research question in literature: '*Does the historical Roman road network play a role on current economy? Is it affected by persistence?*'. The main purpose was to study whether the historical Roman road infrastructure played a role on current trade costs of Italian provinces, analysing how deep and strong the effects of past historical episodes and heritages are in shaping actual economic outcomes and institutions, examining the persistence of the old infrastructure on the present one, and exploring the importance of history in the understanding of the present. On this aim, the analysis has not been limited to the fundamental research on the long-lasting effect of ancient Roman roads on current performance, but it has been equipped with three more pioneering research elements. In fact, beyond the novelty of the Roman roads issue, this chapter adds three more factors of originality. First, it exploits a new measure of Roman roads which captures the length in kilometres of old Roman roads. Second, rather than considering development, this chapter looks at trade and, in particular, at the indirect measure of trade costs constructed according to the 'top-down' approach proposed by Novy in 2013. Third, it applies the indirect measure of trade costs at the sub-country (provincial) level rather than at the country one.

The indirect measure of trade costs computed for the Italian provinces for the period 2003-2010 is the dependent variable of the entire analysis. It is the yearly average of the geometric average of bilateral trade costs between each province and its country partners.

The case of Italy and, more in particular, the NUTS3 level disaggregation is aimed at the cause of identification: Italy is characterised by a historical duality between a developed North-Centre and a less developed South. Moreover, Italian provinces seem to perform differently although the institutions are run by the central government.

Before assessing and exploring the persistence of the Roman road network with classic methods, this chapter has first performed a complete analysis on the potential presence of an endogeneity problem, a typical drawback in the relationship between infrastructure and economic outcomes. On this aim, a detailed investigation on historical sources, narrations and facts has been run in order to understand why and how Roman roads have been constructed. The military reason is the motivation that led to the construction of roads: paved roads facilitated the movement of the army and allowed the transportation of supplies to the troops. The engineering reason is the factor that helped to explain how Roman roads have been built: straight lines between two strategic points. Both military and engineering reasons helped to address the potential endogeneity problem, suggesting that it does not plague the Roman infrastructure and, hence, the relationship between Roman roads and current trade costs. An extra support of the absence of an endogeneity problem has been provided looking at the case of the Via Appia, the most important Roman road, and the conquest of South Italy first and Greece then. Moreover, this chapter went further exploring the relationship between geography, and more in particular the topography of the land, and the Roman road infrastructure.

The empirical analysis proceeded exploiting two different empirical strategies. On the one hand, it focused on the pure and comprehensive effect of the Roman road system, controlling for the past dominations that occurred between the Roman empire and the unification of Italy, the current infrastructure and productivity. In this stage, those mechanisms through which the Roman network performed have been disregarded. To gain more variability, panel data for the dependent variable have been used. Due the time-invariant nature of the Roman road variable, regional, rather than provincial, fixed effects have been included in the model. The check for the current infrastructure has been performed using the IV approach, since the reverse causality issues that may rise and lead to an endogeneity problem. Relying on the literature of Ramcharan (2009), geographical variables have been exploited as instrumental variables. This first empirical strategy has been named 'direct effect', since it has been completely devoted to the effect of the Roman road network leaving out the channels through which Roman roads performed. In the second step of the empirical part, the analysis has been deepened investigating whether the current infrastructure is the channel through which the Roman roads affect present economic outcomes. In light of this, this second empirical strategy has been named 'indirect effect', since it assumes that the Roman road network affect current provincial trade costs through the Italian current infrastructure. Once again, IV methods have been exploited: Roman roads have been adopted as instrumental variable for the current roads. Past dominations and productivity controls, regional and time effects completed the analysis.

Robustness checks have exploited cross-sectional data rather than panel data, have considered the EU15 market rather than the world market and have used lower and higher elasticity of substitution to compute the indirect measure of trade costs.

The evidence coming from the inference analysis confirms expectations and shows a clear and robust negative effect of the integrated ancient Roman road system on current trade costs. The negative effect of the Roman road network is supported in all specifications and in both empirical strategies. The significance and the sign of the Roman road variable coefficient persists when controlling for past dominations, for the current infrastructure and for productivity. The second identification strategy confirms that one channel through which Roman roads perform is represented by current roads. Moreover, robustness checks adopting cross-sectional data, looking at the EU15 market and using a lower and a higher elasticity of substitution for the trade costs measure confirm strongly the significance, the sign and the magnitude of the estimates.

The evidence from the first analysis suggests that provinces with a large Roman road network have a propensity to have more trade relations abroad rather than with themselves. The reasonable key idea behind this is that the Roman road system is affected by persistence. This persistence performed through several advantages and benefits. A denser transportation network enabled more developed and urbanised settlements, more active and functioning cities, more economic activities and trade. This has led to more contacts and relationships between people, shaping a more open mentality and a higher propensity to engage with different peoples and cultures, as it was during the Roman empire. In this perspective, it can be argued that the Roman road system had not only an active part in reducing physical distances, but also a key role in shaping human mind. These conclusions are in line with that strand of the literature that claims the importance of the Roman cultural heritage for the globalised present, due its lasting role in inspiring people and nations (Witcher, 2015), and that asserts how the main inheritance of the Roman world is immaterial, consisting in all those ideas, thoughts and notions that survived until now enclosed in that durable knowledge that comes from the past, without a precise information on when and where they have been acquired (Hingley, 2015).

The instrumental variables method is more inspired to the ʻphysical' rather than to the ʻopenness' concern, proposing how for provinces with a longer and denser Roman network the old infrastructure has represented a good starting point and a basis for the new infrastructure. Further research on this subject will be at the heart of future work.

The evidence produced gives the impression that both ʻphysical' and ʻmental' subjects come along the same conclusion: the Roman empire, with its aim of expansion, development and growth, with its engineering abilities and military capacities, with its well-structured organisation and effective systems, with its culture and advanced knowledge, had such a deep, strong and lasting effect on such a huge variety of concerns, that past facts and old history should not be underestimated and should be considered in providing guidance for policy.

APPENDIX 3.1 - POST-ESTIMATION IV TESTS

According to Baum et al. (2003), an every-present problem in empirical works is the presence of heteroskedasticity. Although IV estimates are consistent and not affected by heteroskedasticity, the standard IV estimates of the standard errors are not consistent, compromising the inference analysis. These problems can be addressed in part through the use of robust standard errors and statistics. Though consistent, the IV estimates are, however, not efficient in the presence of heteroskedasticity. Moreover, the usual forms of the diagnostic tests for endogeneity and overidentifying restrictions will also be invalid if heteroskedasticity is detected. The test of Pagan and Hall (1983) has been designed explicitly for identifying the presence of heteroskedasticity in IV estimation.

Table 3.6 shows the Pagan and Hall's test results. The null hypothesis of homoskedasticity is always rejected. In order to partially deal with heteroskadasticity, clustered standard errors have been included for each specification.

Testing for the failure of the relevance condition is the second check inescapable with IV estimates. Basically, the standard errors in IV estimates tend to be higher than in OLS estimates, and much higher if the excluded instrumental variables are only weakly correlated with the endogenous regressors (first condition not completely or not satisfied). Evidence in favour of the chosen instruments can be found by inspecting the first-stage F-test for the joint significance of the excluded instruments. The first-stage F-test is a “rule-of-thumb” diagnostic according to which, if instruments pass the threshold value of 10, that is the rule of thumb suggested in the literature, they are valid. Table 3.6 reports the first-stage F-statistics: the high values are a proof of the validity of the relevance condition and of the absence of weak instruments for all models. Moreover, Table 3.6 displays the first-stage estimates. Elevation reports a highly significant negative coefficient in all specifications, confirming the negative relationship between elevation and current infrastructure suggested by the literature. Conversely, the coefficient linked to the percentage of mountainous territory is positive rather than negative, suggesting that, the more mountainous the territory is, more kilometres of road exist. The effect is missing for motorways.

The check of the first-stage F-test has confirmed the relevance of selected instruments. The definition of weak instruments has been formalised by Staiger and Stock (1997), who developed the test for the relevance condition. Stock and Yogo (2005) deepen the investigation and formalise Staiger and Stock’s procedure. They provide practical rules of thumb regarding the weakness of instruments, based on the Cragg and Donald (1993) statistic, according to which the tested null hypothesis is that the estimator is weakly identified, in the sense that it is subject to bias that the researcher finds too large.[50] Basically, both Staiger and Stock (1997) and Stock and Yogo (2005) tests reject the null hypothesis of weak instruments when the Cragg and Donald statistic exceeds a specific and determined limit. This test statistic reduces to the first-stage F-statistic in the case with a single endogenous regressor. Moreover, when i.i.d. assumption is dropped and when clustered standard errors are included, the Cragg-Donald-based weak instruments test is no longer valid and the correspondingly-robust Kleibergen-Paap Wald F-statistic should be used (StataCorp, 2013). Although the weak instrument test is suited for more than one endogenous regressor, Table 3.6 shows also, for completeness, the Kleibergen-Paap Wald F-statistic, that is equal to the first stage F-statistic since in all six specifications only one endogenous variable is used. The value of the statistic exceeds the Stock-Yogo critical values of 15% (in two cases 10%) maximal IV size distortion, that is highly acceptable, confirming that none of the instruments used are weak.

It has been mentioned before that IV post-estimation tests could be invalid if heteroskedasticity is detected. The logic of using the first-stage F-statistic relies heavily on the assumption of conditional homokedasticity. Montiel Olea and Pflueger (2013) propose a test for weak instruments that allows for errors that are not conditionally homoskedastic and serially uncorrelated. It extends the Stock and Yogo (2005) weak instrument tests for both 2SLS and Limited Information Maximum Likelihood (LIML) with a single endogenous regressor. Differently from Staiger and Stock (1997) and Stock and Yogo (2005), who test for weak instruments under the assumption of conditionally homoskedastic and serially uncorrelated model errors, Montiel Olea and Pflueger test is robust for heteroskedasticity, autocorrelation, and clustering (Montiel Olea and Pflueger, 2013).[51] The robust weak instrument test rejects the null hypothesis of weak instruments when the effective test F-

---

[50] The null hypothesis of weak instruments can be expressed in terms of estimator bias or test size distortions. Stock and Yogo calculate critical values that are determined by the number of endogenous regressors, the number of instrumental variables, the maximum bias and the estimation procedure. An analogous approach is proposed for the test size, but instead of controlling the bias, it controls the size of a Wald test.

[51] It uses the standard Nagar (1959) procedure to obtain a proxy for the asymptotic estimator bias. The Nagar bias is always defined and bounded for both TSLS and LIML. Montiel Olea and Pflueger define the null hypothesis of weak instruments as if the Nagar bias may be large. Under the alternative hypothesis, the Nagar bias is bounded relative to the benchmark. The benchmark captures the "worst-case" situation when instruments are completely uninformative and when first- and second-stage errors are perfectly correlated. The null hypothesis is that the Nagar bias exceeds a fraction τ of the benchmark for at least some value of the structural parameter and some direction of the first stage coefficients. On the other hand, under the alternative hypothesis, the Nagar bias is at most a fraction τ of the benchmark for any value of the structural parameter and for any direction of the first stage coefficients (Pflueger and Wang, 2015).

statistic[52] exceeds the critical value. This critical value is determined by the significance level, the selected threshold, the estimated variance-covariance matrix and the estimator (2SLS or LIML). Table 3.6 includes Montiel-Pflueger effective F-statistics. The high value of the statistic exceeds the fraction τ=10% for some specifications and 5% for other specifications of the benchmark bias, and allows to reject the null-hypothesis of weak instrument with a 5% significance level.

Testing for the failure of the exclusion condition is the third check inescapable with IV estimates. The Sargan's test, also named test for overidentifying restrictions, tests the null hypothesis that the excluded instruments are valid instruments. According to the exclusion condition, an instrument is valid if it is not correlated with the error term and correctly excluded from the estimated equation.[53] Table 3.6 shows that, when the endogenous regressor is represented by the kilometres of total roads, the exclusion restriction condition is perfectly satisfied. When the instrumented variables are the kilometres of motorways or railroads, the exclusion restriction is no longer valid. The test for overidentifying restrictions is discussed when using panel data.[54]

Last test that needs to be performed is the Durbin-Wu-Hausman (DWH) test for endogeneity of regressors. According to Baum et al. (2003), a Hausman statistic for a test of endogeneity in an IV regression consists in considering the OLS as the efficient estimator and IV as the inefficient but consistent estimator. The test should not be interpreted as a test for the endogeneity or exogeneity of regressors per se, but rather as a test of employing different estimation methods on the same equation. The null-hypothesis is that the residual is zero and that, therefore, the variable being instrumented is exogenous and IV are non needed. As shown in Table 3.6, the DHW does not reject the null-hypothesis in 3 out of 6 specifications, concluding that in those regressions the current infrastructure measure is exogenous and the IV estimator is not required. However, it should be considered that the presence of heteroskedasticity warns about the reliability of tests results.

[52] In the just-identified case with one instrument, the effective F-statistic is equal to the robust F-statistic, but normally it differs from both the non-robust F- and the robust F-statistic.

[53] Statacorp, 2013.

[54] Bowsher (2002).

APPENDIX 3.2 - ROBUSTNESS CHECKS: ESTIMATION TABLES

Table 3.8 Estimation results base model: Roman roads in kilometres and density (panel data, EU15 market, σ=11)

Dependent variable: Average trade costs (ln), σ=11	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	km				**density**			
	Certain	All	Major	Certain & major	Certain	All	Major	Certain & major
Roman roads (ln)	-0.097***	-0.092***	-0.047**	-0.100***	-0.108***	-0.084***	-0.034	-0.113***
Normans	-0.004***	-0.004***	-0.003**	0.001	-0.002*	-0.001	-0.002	0.004**
Swabians	0.001***	0.001*	0.000	0.001**	0.001**	0.000	0.000	0.001**
Anjou	-0.001**	-0.000	-0.000	-0.001***	-0.001***	-0.001***	-0.001*	-0.001***
Spain	-0.000	-0.000	-0.000	-0.000**	-0.000	-0.000	-0.000	-0.000***
Bourbons	0.005***	0.007***	0.007***	0.007***	0.004***	0.006***	0.006***	0.007***
Papal State	0.000	0.000	0.000*	0.000**	0.000	0.000	0.000**	0.001**
Venice	-0.000	-0.000	-0.000	-0.000	-0.000*	-0.000	-0.000	-0.000
Austria	0.000	0.000	0.000	0.000	0.000**	0.000	0.000	0.000**
Savoy	-0.000	-0.000	-0.000	-0.001***	-0.000	-0.000	-0.000	-0.001***
Constant	0.868***	0.969***	0.112	0.229	0.727***	0.772***	-0.117	-0.266
Regional fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Time effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	664	804	736	608	664	804	736	608
R-squared	0.808	0.739	0.711	0.812	0.804	0.721	0.701	0.807

Note: Asterisks denote significance levels; * p<0.10, ** p<0.05 and *** p<0.01.

Table 3.9 Estimation results base model: Roman roads in kilometres and density (cross-sectional data, world market, σ=7)

Dependent variable: Average trade costs (ln), σ=7	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	km				**density**			
	Certain	All	Major	Certain & major	Certain	All	Major	Certain & major
Roman roads (ln)	-0.118***	-0.119***	-0.082**	-0.119***	-0.127***	-0.091	-0.055	-0.126***
Normans	-0.010***	-0.008***	-0.008**	0.000	-0.007***	-0.004*	-0.005*	0.004
Swabians	0.001	0.000	0.000	0.001	0.001	0.000	0.000	0.001
Anjou	-0.001	-0.000	-0.000	-0.001**	-0.002	-0.001	-0.001	-0.001***
Spain	-0.000	-0.000	-0.000	-0.001	-0.000	-0.001	-0.000	-0.001*
Bourbons	0.010***	0.011***	0.011***	0.013***	0.009***	0.009***	0.010***	0.013***
Papal State	0.000	0.000	0.001*	0.001	0.000	0.000	0.001	0.001
Venice	-0.000	-0.000	-0.001	-0.001	-0.001*	-0.001	-0.001	-0.001*
Austria	-0.000	-0.000	-0.000	0.000	0.000	0.000	0.000	0.000
Savoy	-0.000	-0.001	-0.000	-0.002***	-0.001	-0.001	-0.001	-0.002***
Constant	4.006***	4.056***	2.653***	2.754***	3.833***	3.787***	2.260***	2.173***
Regional fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	83	101	92	76	83	101	92	76
R-squared	0.753	0.727	0.686	0.752	0.745	0.705	0.670	0.742

Note: Asterisks denote significance levels; * $p<0.10$, ** $p<0.05$ and *** $p<0.01$.

Table 3.10 Estimation results IV approach for current infrastructure (Roman roads as instruments) and check for productivity (cross-sectional data, world market)

Dependent variable: Average trade costs (ln), σ=8	(1)	(2)	(3)	(4)
Total roads in km (ln)	-0.565***		-0.315**	
Motorways in km (ln)		-0.299***		-0.127**
Total productivity (ln)			-1.106***	-1.476***
Normans	-0.018***	0.010	-0.012***	0.002
Swabians	0.001	0.001	0.000	-0.001
Anjou	0.003***	-0.001**	0.001*	-0.000
Spain	0.000	0.000	0.000	0.000
Bourbons	0.015***	0.080***	0.009***	0.057***
Papal State	-0.000	0.001***	-0.000	0.001**
Venice	0.000	0.000	0.000	0.000
Austria	-0.001*	0.000	-0.001	0.000**
Savoy	0.000	0.000	0.000	0.000
Constant	7.240***	-2.332*	2.099	-5.207***
Regional fixed effects	YES	YES	YES	YES
Observations	83	66	83	66
R-squared	0.647	0.811	0.801	0.892
First-stage F-statistic	5.620	8.110	3.750	2.830
Sargan P-value	0.303	0.107	0.599	0.762
Durbin-Wu-Hausman P-value	0.024	0.221	0.163	0.788
Kleibergen-Paap Wald F-statistic	5.620	8.110	3.750	2.830
Montiel-Pflueger effective F-statistic	6.894	5.977	5.131	2.951
Pagan-Hall P-value	0.954	0.768	0.875	0.236
Instruments	Certain RR km (ln) All RR km (ln)	Certain RR km (ln) Certain & major RR km (ln)	Certain RR km (ln) All RR km (ln)	Certain RR km (ln) Certain & major RR km (ln)

Note: Asterisks denote significance levels; * p<0.10, ** p<0.05 and *** p<0.01.

Chapter 4

COUNTRY TRADE COSTS: DIFFERENT DETERMINANTS BY DIFFERENT GEOGRAPHIES?

4.1 Introduction

Trade costs are like a box including all those elements that separate countries and reduce the probability that trade between two partners takes place. Although literature on trade costs has reduced the unawareness about their components, the research on the extent and the importance of these elements in determining trade costs is poor. Due data constraints, identifying what affects trade costs for a defined country or pair is hard, making trade costs a black box at the disaggregated level. Nevertheless, analyses on groups of countries are more feasible, and new research has started to decompose trade costs into their components with the precise aim to address policy-makers in taking the best strategies for reducing the trade wedge between countries.

The aim of this chapter is completely devoted to understanding why and what makes trade between countries costly. Starting from a set of different potential and more likely determinants of trade costs, the intention is to identify, what are the most influential sources able to explain the extent of trade costs between countries. The inclusion of a wide set of potential determinants of trade costs represents a challenge in this kind of analysis, since the problems of data availability.

The indirect measure of trade costs, computed according to the ‘top-down’ approach proposed by Novy (2013), is the dependent variable of the entire analysis. It represents an ideal measure to capture the trade costs issue. In fact, the way according to which indirect trade costs are constructed allows, on the one hand, to assess the participation of countries to trade and, on the other hand, to measure how large trade costs are. And these two facts are strictly connected. The logic behind the indirect measure of trade costs, as widely exposed in Chapter 1, is that, when international trade costs between countries i and j, relative to domestic trade costs within each country, are low, then countries trade more internationally than domestically. In this framework, having low trade costs determines more the participation to international trade markets rather than to the domestic ones, and helps to explain why some countries take part to international exchanges and others don't or why for some countries it is easier to trade internationally than internally.

It should be kept in mind that the indirect measure of trade costs is a symmetric measure of bilateral trade costs, since it is the product of the trade flows of two trading partners. In this light, the case of zero trade flows acquires a particular importance in the context of trade costs: if discarded, then trade costs will be missing ('unadjusted' version of trade costs); if transformed in ones, then trade costs will be positive ('adjusted' version of trade costs).[1] The case of missing trade costs represents a further form that trade costs may assume (positive, zero, negative, missing).

Both versions of trade costs have pros and cons. This has led literature to exploit both measures, selecting the one closer to the purposes of the analysis carried out. This chapter performs similarly, but meanwhile it takes advantage of the puzzle, adopting the non-selected measure for robustness checks.

The study of why and what makes trade between countries costly starts from the analysis of zero trade flows and proceeds with the investigation of the potential factors affecting trade costs. This two-stage analysis is performed relying on the Heckman (1979) model. In a first stage, the investigation of zero trade flows helps to consider the case of missing trade.[2] In a second stage, the extent of trade costs is assessed looking at what makes trade costly and exploring the main dimensions that are linked to trade. For this kind of analysis, the 'unadjusted' version is more appropriate.

This double and structured approach of studying trade costs is enhanced with the geographical perspective. Differently from previous works, the interest is focused on geographies of countries: rather than simply analysing what are the best determinants of trade costs for a group of countries in the world, here the aim is to assess whether the determinants vary

[1] Missing trade costs may occur not only when trade flows are zero, but also when the information about GDP is missing or when internal trade is negative (GDP < total exports).

[2] Literature on trade defines trade as absent when, given a country pair i-j, i does not export to j and j does not export to i. In this chapter, also intermediate cases are considered: i exports to j but j does not export to i, or i does not export to j but j does.

among different geographies and, if so, which is the extent of each determinant for each geographical category. On this aim, 188 world countries are distinguished in four different groups according to their geography: landlocked countries, coastal countries, partial-insularity countries, island countries.

The insight of classifying countries in different geographical groups derives from the fact that geography matters. It matters because it influences development and the abilities of countries to trade without barriers, to participate to international markets and to promote sustainable growth. And the motivation of considering geographical categories as 'degrees of insularity' is given by the very recent findings of the literature. Whereas the conditions of being landlocked or of being a small and remote island in the ocean[3] have been highly recognised as being the worst immediate cases of 'bad' geography,[4] the literature has not equally considered the intermediate conditions of insularity. Pinna and Licio (2013) attempt to fill this gap constructing a new insularity data set and a new measure of insularity for all countries in the world, able to capture for each country its 'degree of insularity'. The categorisation of the 188 countries in the sample originates from the Pinna and Licio's (2013)[5] contribution.

The closer contributions in literature to this chapter are represented by the works of Arvis et al. (2013a, 2013b) and Duval and Utoktham (2011b). Arvis et al. (2013a)[6] find that the distance between countries has the main influence on trade costs. Maritime connectivity and logistics abilities account for a significant share, and, when considered together, their effect is equivalent to the one of the physical distance. Duval and Utoktham (2011b) are more oriented to the effects of trade-facilitation and policy variables for the ASEAN developing countries, suggesting that port services, logistics, communication facilities and access to information play a fundamental role in determining trade costs and that ad-hoc policies may mitigate international trade costs.

In terms of originality, this chapter tries to contribute to the literature on the determinants of trade costs in three ways. First, a large number of countries distinguished by geographies allows to disentangle the effects of each determinant according to the 'degree of insularity', providing also further evidence to the literature on the importance of geography in affecting economic outcomes. Second, the focus on trade costs in conjunction with a two-steps

[3] This is the case of the Small Islands Developing States (SIDS).

[4] See, among others, Briguglio (1995) for a discussion on the SIDS and Arvis et al. (2010) for the landlocked countries.

[5] See Subparagraph 4.3.1 for the definition of each geographical (insular) category and to know how categories have been treated.

[6] Arvis et al. (2013a) infer trade costs for up to 178 countries over the period 1995-2010, with a special focus on developing countries.

estimation procedure based on the Heckman (1979) model is particularly suitable for the 'unadjusted' version. The fact of being not adjusted for zero trade flows allows, in a first step of the Heckman approach, to assess the probability of positive trade and to explore the extensive margin of trade. In the second step, it is mainly the intensive margin the core of the investigation: positive trade costs are examined to explore what are the main and potential determinants affecting them. Third point of innovation is represented by the discussion of the Bayesian Model Averaging (BMA) related to the trade costs analysis. The BMA approach is a novelty in the trade costs field and the idea, that will be developed in future research, is basically to import in the analysis of the determinants of trade costs, characterised by model uncertainty, what Crespo Cuaresma et al. (2009) perform for economic growth.

This chapter consists of eight paragraphs. In Paragraph 4.2, the central role of geography and natural barriers in determining an economic disadvantage in terms of development and trade is examined. The intent is to understand and appreciate how geography may represent a critical obstacle in creating connectivity with trading partners and in fostering economic growth. In Paragraph 4.3, a twofold descriptive analysis is provided. On the one hand, main figures about trade costs are presented, in order to examine the case of missing trade costs and, therefore, the case of zero trade flows, and to explore what dissimilarities occur between countries with different geographies, what has been the pattern during last twenty years and what are the differences between rich and developing countries. On the other hand, data sources and descriptives are provided for the main determinants of trade costs on the identical aim to appreciate variations across geographies. Paragraph 4.4 presents the empirical model. The results obtained from the empirical analysis are discussed in Paragraph 4.5. Paragraph 4.6 includes all robustness checks performed. Paragraph 4.7 discusses about the best approach able to capture and measure the determinants of trade costs. An overview of the Bayesian Model Averaging method related to the case of trade costs is presented as a possible new path to follow with future research. In Paragraph 4.8, some concluding remarks are provided. The chapter ends with Appendix 4.1, which includes all tables referring to the robustness checks.

4.2 Geographical disadvantage and trade

The importance of geography in explaining the wealth of countries has been highly underlined by the economic research. Constraints given by nature are responsible for countries' isolation and openness and determine participation in international markets. The persistent spirit of geography is invoked as the original cause of underdevelopment of countries and as the root of poor institutions, education and social capital. Moreover, countries subject to geographical barriers, experience a supplementary disadvantage in terms of attracting skilled workers, physical capital and new knowledge, compromising current and future development.

The attention on the effect of geography on development began in modern times with Gallup, Sachs and Mellinger (1999); starting from then, insights from the new economic geography, trade theory and growth literature have been applied to stress the link between geographic features, trade and national wealth. However, since classical economics, the interest on how geography correlates with growth has been one of the issues of the ancient literature. Adam Smith (1776) argues that the role of geography is fundamental for transport costs and, in turn, for productivity.

Talking about geography is not an easy task. Geography is a wide topic that may lead the narrator to open several doors to all those themes that are directly or indirectly linked with the geographical issue. Geography has to do with institutions, with infrastructures, with urbanisation; it is connected with poverty, health, education. Basically, geography is strictly connected with the fortune of countries. In order to address the geographic theme coherently to the purposes of this chapter, two main areas of discussion have been identified. The first area discusses the key role of geography for economic development of countries, examines what are the channels through which geography is able to affect the performance of nations and explores those geographical features that are related to the 'productivity ease' (climate, latitude, terrain ruggedness) or to the 'connectivity-openness' (being landlocked, having access to sea) of countries. The second area takes, instead, the trade costs perspective. On these bases, following Subparagraph 4.2.1 looks at the first area, presenting a very brief review of the main findings coming from the literature. Subparagraph 4.2.2 focuses on the second area and examines the relationship between trade costs and geography.

4.2.1 The importance of geography and natural features

The three main sources of economic development are represented by health environment, institutions and market integration. At the basis of these factors there is geography. Geography and climate can be considered the effectively exogenous factors of growth; health, institutions and trade openness are endogenous. As underlined in Chapter 2, the strand of economic literature, that looks at the long-terms determinants of development, divides into that supporting the '*institutional approach*' and that which is more prone to confirm the '*geographical approach*'. According to the '*pro-institutions*' view, institutions are the only ones determining development and growth in the long-run; geography has an indirect effect through its influence on institutions. The '*pro-geography*' vision rejects this stringent perspective and argues that, although institutions are certainly the main factor explaining why some countries are trapped by an underdeveloped and poor environment, they are not able to explain everything.

According to Sachs (2001), there are four main channels through which geography affects the performance of countries: health, agricultural productivity, physical location, natural resources availability. All these factors are strictly connected with biological variables, like climate, temperature and precipitation (Dell et al., 2012; Sachs, 2001; Kamarck, 1976; Myrdal, 1968), terrain conformation (Nunn and Puga, 2012; Riley et al., 1999) and the spatial connection in terms of having access to the sea (Rappaport and Sachs 2003;), being landlocked (Arvis et al., 2010; Faye et al., 2004) or being an island (Becker, 2012; Borgatti, 2008; Mimura et al., 2007; Briguglio, 1995; Briguglio and Kaminarides, 1993).

The important role played by climate and latitude in affecting the economic development of countries is corroborated since Montesquieu (1750). Over the years, several contributions have provided new insights on the relationship between climate/latitude and economic performance, focusing on different facets (agriculture, migration, tourism, mortality, crime, conflicts, etc.). Some historians, like Jones (1981) and Crosby (1986), retrieve in geography and climate a valid explanation for the economic development of European peoples and their success in the conquest and colonisation of lands over the world, corroborating what some years later will be argued by Diamond (1997). Diamond focuses more on health. The health theme and its important role in affecting the development of economies has been extensively discussed in literature. Safe environment and sanitary conditions enhance productivity via healthy workers. According to Diamond (1997), the success of European civilisation and its ability to expand throughout three continents (America, Africa, Oceania) was not due to an intellectual superiority of Europeans, rather it was because they had the fortune to live in a continent (Eurasia) whose environmental conditions favoured the development and the diffusion of two elements that contributed significantly to the European supremacy on other peoples: weapons and diseases. The high population density of the Eurasian cities and their network of connections with other cities led inhabitants of Eurasia to develop a partial immunity against highly contagious diseases. In other continents, since the lack of comparable societies and of complex environmental conditions, a similar immunity has not been developed by the peoples: 90 percent of Native Americans were killed by diseases introduced by Europeans during the conquest of Americas. This historical view is validated also by other scholars. Olsson and Hibbs (2005), Easterly and Levine (2003), but also the contributions of Engerman and Sokoloff (1997, 2002) and Acemoglu, Johnson and Robinson (2001, 2002), widely discussed in Chapter 2, retrieve in history the channel through which geography (and in particular the latitude and the climate) affects development.

Dell et al. (2012) focus on temperature and precipitation. Measuring the temperature and the precipitation for all countries in the world and for a time span of more than fifty years, they find that the disadvantage that originates from higher temperatures involves only poor countries, where both level of output and growth rates are affected. In more developed countries, the effect is little. Temperatures perform mainly via agricultural productivity, but industrial output, investments and institutions are also important channels through which temperatures affect economic growth. Differently, precipitations have not such an impact on performance and development. On the same view, Sachs (2001) underlines how economies with tropical climate are mainly in a poor economic condition, whereas countries in temperate eco-zones benefit from a wealthier state. The grounds for underdevelopment of tropical regions lie, according to Sachs, in three main elements deeply rooted in the relationship between physical ecology and development: technology, urbanisation and institutions. In tropical eco-zones, ecosystem, genetic features of peoples, malnutrition, poverty and the lack of medical care services favour and have favoured in ancient times the transmission of diseases and epidemics more than in other eco-zones. In such environments, to control for the diffusion of illnesses and to eradicate viruses and bacteria responsible for infections becomes more complicated than in other zones. Analogously, in tropical regions agricultural productivity is lower than in temperate zones, due a set of physical elements, like high temperature, water availability, fertility and erosion of the soil, parasites and pests. Agricultural progresses, increasing the productivity of terrains and reducing infections of plants, are arduous in similar natural systems. Similarly, the availability of energy sources and the possession of deposits, like oil, coal or hydrocarbons, is fundamental to perform advances in manufacturing productions. If natural conditions are crucial in allowing the development and the diffusion of technological innovations able to produce increasing returns to scale from production activities, geographical features play also an important role in determining population localisations and institutions, other two key elements for the economic growth of countries (Sachs, 2001).

Nunn and Puga in 2012 focus on another trait of the land: ruggedness. According to the authors, ruggedness represents a geographical disadvantage in economic terms, since it implies costs and defies. A rugged territory is difficult to farm: arable lands are few. A rugged territory is difficult to cross: transportation infrastructures, as examined in Chapter 3, are scarce and time and costs of transport are higher. A rugged territory is inhospitable: population prefers to locate along coasts, where water sources are available and connectedness is easier. The contribution of Nunn and Puga (2012) is important. While, on the one hand, they stress the negative effect on economic outcomes, due weaknesses in trade and agricultural productivity, on the other hand, exploiting a specific terrain ruggedness index, they find that in Africa having a rugged territory represented a benefit in the past. The reason should be retrieved in history. Rugged territories served as protection from slaves traders between 1400 and 1900 and, hence, from subsequent late economic development: slaves exports were one of the reasons of the underdevelopment of African countries. However, today geographical ruggedness in Africa, as in the world, represents an economic drawback for performance and growth. Nevertheless, Nunn and Puga show that for African countries the positive historical (indirect) effect of ruggedness is not overtaken by the negative current (direct) effect, confirming the long-term effects of geography via history and corroborating the historical approach.

Some lines above, it has been highlighted how countries located in tropical zones are essentially poor, while countries with temperate climate are more economically developed.[7] This insight coming from Sachs (2001) is not complete. Indeed, Sachs (2001) argues that, when economies are not developed in temperate zones, it is because they were under the communist rule or they simply are geographically isolated. Isolation is one of the main issues in the geographical disadvantage theme. For islands, size and remoteness are the critical factors that impede development. For landlocked countries, it is, instead, the lack of access to the sea. Sachs (2001) observes that coastal countries are generally richer than landlocked countries. Accordingly, Bloom and Williamson (1998) suggest that nations having a coast benefit from the most advantageous geographic condition for performance and development. Arvis et al. (2010), in the World Bank contribution "The cost of being landlocked", highlight how Africa, where one country out of three is landlocked, faces more than other continents or macro-regions the economic disadvantage of being landlocked: the fraction of landlocked countries in the world is one out of four. This economic disadvantage performs in higher transaction costs. In Africa, the transport costs are the highest in the world. Islands face a similar drawback. Although some literature, like Armstrong et al. (1998) and Bertram and

[7] Using data from Olsson and Hibbs (2005), Spolaore and Wacziarg (2013) show how a set of geographical variables, that account for latitude, tropical climate, landlocked or island territory, jointly describe 44 percent of variation in modern log per capita income, and that latitude explains the largest effect.

Karagedikli (2004), is more prone to highlight the economic gains deriving from the insular condition, like tourism, ease in social capital building or in policy implementation, other scholars, like Briguglio (1995) and Dommen and Hein (1985), argue how islands are particularly exposed to critical economic difficulties, due remoteness, smallness and vulnerability. These economic difficulties perform in a reduced international trade, in a small domestic market, in limited resources, in higher migration.

This paragraph ends with the seminal contribution of Gallup, Sachs and Mellinger (1999), the work that marked the rebirth of interest of economic literature in geography. Gallup and co-authors put together different issues linked to the geographical sphere and provide a widespread perspective on how natural features affect economic performance and development. First, tropical regions, rather than temperate eco-zones, face the main limitations in economic growth because of diseases (like malaria) and reduced agricultural productivity, that typically are connected with the tropical climate. Second, having access to the sea represents an opportunity and a benefit in development. Coastal regions, rather than landlocked ones, experience the main advantages of their location. The access to the coast reduces transport costs and generates agglomeration economies. Countries having a coast can exploit the gain coming from coastal trade and benefit from increasing returns to scale in labour. Third, population density is higher in temperate zones, in lands with good soil and in regions having access to water. It is more the proximity to inland or ocean-navigable rivers, rather than the proximity to coast, that determines the settlement of people. High population density has a positive effect on growth in coastal regions, where increasing returns to scale from transportation infrastructure and the division of labour may advance domestic and international trade. In inner regions this is less the case. However, in modern times, growth in population is negatively correlated with economic growth (Gallup et al., 1999).

4.2.2 Geography and trade costs

The brief discussion on the relationship between geography and trade costs starts from the seminal contribution of Anderson and van Wincoop (2004) on trade costs, extensively discussed in Chapter 1. Anderson and van Wincoop (2004) underline how the internal geography of countries is fundamental to understand trade costs and resulting international trade volumes.

A large part of the literature has focused on the important role played by the distance on trade costs. The contribution of Chen and Novy (2011) is one of the most representative. Chen and Novy, by developing a micro-founded measure that captures bilateral trade frictions across countries and industries, use, among others, a set of geographical variables (international distance, domestic distance, adjacency, common language) to assess what factors affect trade integration. They find that relative trade costs increase with international distance and decrease with domestic distance, but the effect of the latter one is almost twice the first one. Having a common border positively affects trade integration: between adjacent countries trade costs are 7 percent lower than between nations that do not share a border. Speaking the same language also has a positive effect on trade integration, yielding 14 percent lower trade frictions.

Beyond the mere geographical distance, trade costs literature has also argued that distance is not the only geographical factor important in determining trade costs. Being landlocked, being remote or being an island have the effect of making trade more costly. On these bases, Behar and Venables (2010) highlight how countries that don't have access to the sea experience trade costs that are 50 percent higher than in other countries. The component of trade costs that particularly faces the burden of the geographical disadvantage is represented by the transport cost. Limão and Venables (2001) show that low quality infrastructure and unfavourable geographical conditions, like being completely surrounded by other countries, make transport costs higher. Behar and Venables (2010) confirm this idea, arguing that part of the extent of transport costs is due to geographical distance and lack of a coast. Korinek (2008) looks at maritime transport costs, a specific component of transports costs, and in turn of trade costs, and argues how for remote and small countries, most of them islands (Guam, Nauru, Christmas Island, Tonga, Pitcairn) or African states (Guinea, Sierra Leone, Togo), maritime transport costs represent a significant obstacle to trade. She suggests that, due these high costs, these countries should specialise in producing, and therefore exporting, goods for which transport costs are low and having, therefore, a very high value to weight ratios.

Although it has been highly remarked the importance of geography in determining trade costs and how adverse geography represents a concrete burden for those countries where geography can be defined 'bad', research in the trade costs field is still needed and should go beyond the mere physical distance. This chapter attempts to provide empirical evidence in this sense.

4.3 Data and descriptives

To understand what are the main determinants of trade costs, whether these determinants vary across geographies of countries, and, if so, what is the extent of these determinants for each geographical category, is the main aim of the chapter. As underlined above, the analysis on the determinants of trade costs presents some difficulties in terms of data availability and limited observations when applying econometric methods. Moreover, the study of the main determinants of trade costs cannot escape from a deep understanding of the trade costs issue. The case of missing trade and the size of trade costs by geographies of countries should be deeply analysed in order to value trade costs in both extensive and intensive margins. Therefore, before dealing with the empirical model, a complete descriptive analysis based on figures and statistics is required to understand what is the conduct of the data and to appreciate patterns and potential phenomena pictures. On this aim, this paragraph has been divided into two parts. First, all descriptives on trade costs are presented in order to take into account both zero trade flows and, in turn, missing trade costs and the extent of trade costs. Since zero trade flows typically imply prohibitive fixed costs,[8] the ‘unadjusted’ indirect measure of trade costs interprets the case of missing trade costs as non-participation to international exchanges. Second, all descriptives on the main factors potentially affecting trade costs (the independent variables) are produced, to have a picture of the main spheres connected to the trade costs issue.

4.3.1. Trade costs

Trade costs have been computed according to the methodology proposed by Novy (2013)[9] and further examined in Jacks et al. (2008), Chen and Novy (2011) and Chen and Novy (2012). As examined in Chapter 1, the measure is basically a comprehensive aggregate measure of bilateral trade costs, including all possible costs connected with international trade (transport costs, tariff and nontariff measures, distance, institutions, language and culture, trade facilitation bottlenecks, etc.). It originates from the insight of the structural gravity model by Anderson and van Wincoop (2003) and the additional contributions coming from Anderson and Yotov (2010, 2012), Fally (2015), Head and Mayer (2014) and De Benedictis and Taglioni (2011). The equation of the bilateral flow between country i and country j, expressed according to Anderson and van Wincoop (2003), is enhanced by domestic trade, according to Novy (2013), Jacks et al. (2008) and Chen and Novy (2011), obtaining a geometric average of bilateral international trade costs between country i and country j, relative to domestic trade costs within each country. Differently from Chapter 3, where trade

[8] Novy (2013).

[9] Novy began working on the indirect measure of trade costs some years before 2013. 2013 refers to the date when the contribution “Gravity redux: measuring international trade costs with panel data” has been published on Economic Inquiry.

costs are expressed as means by province, here the more traditional bilateral structure of trade costs at the country level is exploited.

The selected measure for the main empirical analysis is the ʽunadjusted’ version of trade costs at the country pair i-j dimension. Differences between ʽadjusted’ and ʽunadjusted’ versions perform in the different interpretations they entail. This point is particularly important to better appreciate the information provided by the ʽunadjusted’ version in following descriptive and empirical analyses. The ʽadjusted’ version of the indirect measure of trade costs allows to avoid the sample selection bias that arises with zero trade flows and missing trade costs, and performs absent exports as infinitesimal exports that originate from elevated trade costs. The insight coming from this ʽadjusted’ version is in some way related to both intensive and extensive margins of trade: high trade costs lead to lower trade and, therefore, to lower quantities traded (intensive margin) or to lower trade relationships (extensive margin). However, the approach of replacing zeros in the data set with one, as for the ʽadjusted’ measure, does not allow to visibly consider whether the fixed cost to enter in foreign markets is so high that it completely impedes trade or just limits it, yielding tiny exports. On the other hand, the ʽunadjusted’ trade costs cannot be measured when trade is absent in both directions or in cases of asymmetric trade (trade just in one direction), reducing the observed sample. Nevertheless, the fact of being missing when trade is zero allows to easily interpret when trade is absent, and high values of the ʽunadjusted’ measure may be read in terms of both extensive and intensive margins.

Analogously to Chapter 3, trade costs measures have been computed using $\sigma=8$, according to the evidence coming from the most relevant contributions to this approach. Whereas the entire descriptive analysis refers to this estimation of the elasticity of substitution, the empirical analysis[10] exploits lower ($\sigma=7$) and higher ($\sigma=11$) estimates.

Data sources refer to BACI-CEPII data set for bilateral data on trade by country i-country j and total exports by country, and to the World Bank WDI database for the information about GDP. Trade costs are computed for 188 countries. Countries may be distinguished in four geographical categories: 32 landlocked countries, 87 coastal countries, 17 partial-insularity countries, 52 island-states. The classification comes from the Pinna and Licio's (2013) insularity data set. The work, collecting available physical geographic data on islands, calculates a measure of insularity for more than 200 countries in the world, in order to evaluate whether and when insularity represents a ‘bad’ geography condition and when it

[10] Robustness checks.

actually appears to be associated with better economic performance at the country level. The measure is able to capture the heterogeneity associated with the insular condition. This chapter does not use the measure of insularity *per se*, but it considers countries according to their degrees of insularity, that derive in turn from the measure of insularity. The geographical categories work like dichotomous variables (i.e. 1 if the country has a partial insularity, 0 otherwise). Pinna and Licio (2013) define the degrees of insularity using five categories:

- landlocked: countries that don't have access to the sea;
- coastal-zero: countries that have coast but not islands;
- coastal-negligible: countries that have coast and islands, but the percentage of insular territory, with respect to the total country territory, is less than 2 percent;
- partial-insularity: countries that have coast and islands, and the percentage of insular territory, with respect to the total territory, is equal or higher than 2 percent;
- islands-states: countries that are completely surrounded by the sea or the ocean.

The descriptive and the inference analysis of this chapter consider four rather than five categories: countries included in 'coastal-zero' and 'coastal-negligible' categories have been collected in a single category named 'coastal'. Table 4.1 lists all 188 countries for which the information about trade costs is included in the data, according to their geographical (insular) category.

In order to provide the best descriptive of trade costs, in addition to the bilateral dimension country pair i-j, trade costs have been computed as geometric means at the country and world level, exploiting other four different dimensions: time, world, geographical category and WB income category.[11] First, trade costs are investigated for a long 20 years period (1995-2014).[12] Second, trade costs are explored at the aggregate-world level, simply computing the yearly average of the initial bilateral measure. Third, trade costs are studied at the geographical level of country i. Fourth, trade costs are examined at the WB income level of country i. These four dimensions allow to better appreciate regularities and patterns in terms of sample composition, trend across years, variability among geographies of countries and income categories and to provide an overview at the global level.[13]

The data set is composed by 351,560 bilateral flows from 1995 to 2014 for 188 countries in the world. Each pair is included once. Table 4.2 gives an overview of the data set in terms of

---

[11] The information about the WB income group is updated for each year. Hence, as in the definition of the World Bank, countries may change category in the data set.
[12] The inference refers to a more limited time span.
[13] The empirical analysis exploits uniquely trade costs at the country pair i-j level. Means by country are not adopted and serve for the sole descriptive investigation.

Table 4.1 Geographical category ('degree of insularity') by country

Country	Geographical category	Country	Geographical category	Country	Geographical category	Country	Geographical category
Afghanistan	Landlocked	Denmark	Partial	Kuwait	Partial	Rwanda	Landlocked
Albania	Coastal	Djibouti	Coastal	Kyrgyzstan	Landlocked	S. Kitts & Nevis	Island
Algeria	Coastal	Dominica	Island	Lao People's D.R.	Landlocked	S. Vincent & the G.	Island
Angola	Coastal	Dominican Republic	Island	Latvia	Coastal	Saint Lucia	Island
Antigua & Barbuda	Island	East Timor	Island	Lebanon	Coastal	Samoa	Island
Argentina	Coastal	Ecuador	Partial	Liberia	Coastal	San Marino	Coastal
Armenia	Landlocked	Egypt	Coastal	Libya	Coastal	Sao Tome & P.	Island
Aruba	Island	El Salvador	Coastal	Lithuania	Coastal	Saudi Arabia	Coastal
Australia	Island	Equatorial Guinea	Partial	Macau	Coastal	Senegal	Coastal
Austria	Landlocked	Eritrea	Coastal	Macedonia	Landlocked	Seychelles	Island
Azerbaijan	Landlocked	Estonia	Partial	Madagascar	Island	Sierra leone	Coastal
Bahamas	Island	Ethiopia	Landlocked	Malawi	Landlocked	Singapore	Island
Bahrain	Island	Fiji	Island	Malaysia	Partial	Slovak Republic	Landlocked
Bangladesh	Coastal	Finland	Coastal	Maldives	Island	Slovenia	Coastal
Barbados	Island	France	Partial	Mali	Landlocked	Solomon Islands	Island
Belarus	Coastal	French Polynesia	Island	Malta	Island	Somalia	Coastal
Belgium-Lux.	Coastal	Gabon	Coastal	Marshall Islands	Island	South Africa	Coastal
Belize	Coastal	Gambia	Coastal	Mauritania	Coastal	Spain	Partial
Benin	Coastal	Georgia	Coastal	Mauritius	Island	Sri Lanka	Island
Bermuda	Island	Germany	Coastal	Mexico	Coastal	Sudan	Coastal
Bhutan	Landlocked	Ghana	Coastal	Micronesia	Island	Suriname	Coastal
Bolivia	Landlocked	Greece	Partial	Moldova	Coastal	Sweden	Coastal
Bosnia-Herzeg.	Coastal	Greenland	Island	Mongolia	Landlocked	Switzerland	Landlocked
Brazil	Coastal	Grenada	Island	Morocco	Coastal	Syria	Coastal
Brunei D.	Island	Guatemala	Coastal	Mozambique	Coastal	Tajikistan	Landlocked
Bulgaria	Coastal	Guinea	Coastal	Myanmar	Coastal	Tanzania	Coastal
Burkina Faso	Landlocked	Guinea-Bissau	Partial	Nepal	Landlocked	Thailand	Coastal
Burundi	Landlocked	Guyana	Coastal	Netherlands	Coastal	Togo	Coastal
Cambodia	Coastal	Haiti	Island	New Caledonia	Island	Tonga	Island
Cameroon	Coastal	Honduras	Coastal	New Zealand	Island	Trinidad & Tobago	Island
Canada	Partial	Hong Kong	Partial	Nicaragua	Coastal	Tunisia	Coastal
Cape Verde	Island	Hungary	Landlocked	Niger	Landlocked	Turkey	Coastal
Cayman Islands	Island	Iceland	Island	Nigeria	Coastal	Turkmenistan	Landlocked
Central African R..	Landlocked	India	Coastal	Norway	Partial	Tuvalu	Island
Chad	Landlocked	Indonesia	Island	Oman	Coastal	U.S. of America	Coastal
Chile	Partial	Iran	Coastal	Pakistan	Coastal	Uganda	Landlocked
China	Coastal	Iraq	Coastal	Palau	Island	Ukraine	Coastal
Colombia	Coastal	Ireland	Island	Panama	Coastal	U.A.E.	Coastal
Comoros	Island	Israel	Coastal	Papua New Guinea	Island	United Kingdom	Island
Congo	Coastal	Italy	Partial	Paraguay	Landlocked	Uruguay	Coastal
Congo, Dem. Rep.	Coastal	Jamaica	Island	Peru	Coastal	Uzbekistan	Landlocked
Costa Rica	Coastal	Japan	Island	Philippines	Island	Vanuatu	Island
Cote d'Ivoire	Coastal	Jordan	Coastal	Poland	Coastal	Venezuela	Coastal
Croatia	Partial	Kazakhstan	Landlocked	Portugal	Partial	Vietnam	Coastal
Cuba	Island	Kenya	Coastal	Qatar	Coastal	Yemen	Coastal
Cyprus	Island	Kiribati	Island	Romania	Coastal	Zambia	Landlocked
Czech Republic	Landlocked	Korea (South)	Coastal	Russia	Coastal	Zimbabwe	Landlocked

Source: Author's elaborations from Pinna and Licio (2013) insularity data set

number of countries, number of observations, percentage composition and WB income category by geographical group (i.e. 'degree of insularity'). Not all countries have the information about trade costs for the whole period 1995-2014, but each country, out of the total 188, includes the information for at least one year. Table 4.2 allows to appreciate for

Table 4.2 Complete and reduced data set description

	Whole sample	Landlocked	Coastal	Partial-insularity	Island-states
Number of countries	188	32	87	17	52
Number of observations Complete data set	351,560	57,240 (i) 62,440 (j)	155,240 (i) 170,140 (j)	36,860 (i) 26,720 (j)	102,220 (i) 92,260 (j)
Number of observations With info about trade costs	172,660	25,065 (i) 27,163 (j)	83,209 (i) 95,940 (j)	25,462 (i) 18,214 (j)	38,924 (i) 31,343 (j)
Percentage composition Complete data set	100%	17%	46%	9%	28%
Percentage composition With info about trade costs	100%	15%	52%	13%	20%
High income	26%	10%	17%	67%	38%
Upper-middle income	20%	10%	22%	20%	22%
Lower-middle income	27%	21%	32%	6%	30%
Low income	27%	59%	29%	7%	10%

Source: Author's elaborations

both complete data set and reduced data set (i.e. bilateral flows for which the information about trade costs is not missing) the representation of country partners (i and j) in terms of 'degree of insularity'. 46 percent of the sample is composed by coastal countries, 28 percent are islands-states, 17 percent landlocked countries and 9 percent partial-insularity countries. These percentages are obtained as means, considering the size of both partner i and j groups. It is interesting to examine how, looking at the reduced data set (i.e. bilateral flows having the information about trade costs), these percentages vary: coastal partners, involved in 46 percent of total bilateral trade, represent 52 percent of flows having the information about trade costs; island-states, that represent 28 percent of total bilateral flows, account for 20 percent in the reduced data set. The representation of the other two geographical categories do not vary in considerable terms between complete and reduced sample. Another important information is the one provided by the WB income category.[14] The bottom part of Table 4.2 examines how the 188 countries divide into WB income categories, in order to assess whether the four income level categories are balanced within the total sample and within each geographical group. The whole sample presents an adequate equilibrium in terms of income

[14] The World Bank provides an analytical classification of the world's economies based on estimates of gross national income (GNI) per capita for the previous year. According to its GNI, each economy is classified as high income economy, upper-middle income economy, lower-middle income economy, low income economy. This classification is revised every year and countries may change category from one year to the next.

representation: each category accounts roughly for one fourth, even though the upper-middle income category represents the 20 percent of the entire sample. When looking at the landlocked group, the balance between income categories is far from being reached: 60 percent of landlocked countries are poor economies, 80 percent when grouping low and lower-middle income categories together. These figures are not unexpected, since the majority of nations that don't have access to the sea are African states. Within the coastal group, once more, 60 percent of countries are in poor condition, even though the distribution in the four WB income categories is more balanced. When considering the partial-insularity countries, almost 90 percent of the group is composed by rich or middle-rich economies. The equilibrium improves for the island-states group: 60 percent of countries belong to high or middle-high categories, 40 percent are in the lowest categories.

Table 4.3 Matrix of bilateral flows by geographical category: number of observations and percent weight in data set

i \ j	Landlocked	Coastal	Partial-insularity	Island-states	Total
Landlocked	9920 3%	27640 8%	4520 1%	15160 4%	57240 16%
Coastal	28040 8%	74820 21%	11620 3%	40760 12%	155240 44%
Partial-insularity	6360 2%	17960 5%	2720 1%	9820 3%	36860 11%
Island-states	18120 5%	49720 14%	7860 2%	26520 8%	102220 29%
Total	62440 18%	170140 48%	26720 7%	92260 27%	351560 100%

Source: Author's elaborations from BACI-CEPII data, from the World Bank WDI database and from Pinna and Licio (2013) insularity data set

The information provided by Table 4.2 is enhanced with Table 4.3, that displays the matrix of all bilateral flows in the data set with the equivalent percent weight and allows to appreciate the involvement of countries by geographical category. Since the highest fraction of countries is represented by coastal states, the majority of bilateral flows involve coastal countries as trading partners. Values included in Table 4.3 are not interesting *per se*, but they become interesting when considering how many bilateral flows between countries are represented by zeros. In fact, the number of bilateral flows between countries becomes attractive when unravelling which fraction of those flows is represented by positive trade flows in both directions, by zero trade flows in both directions and by asymmetric trade flows (zero flows in one direction and positive flows in the other). On this purpose, Table 4.4 reports, for each of

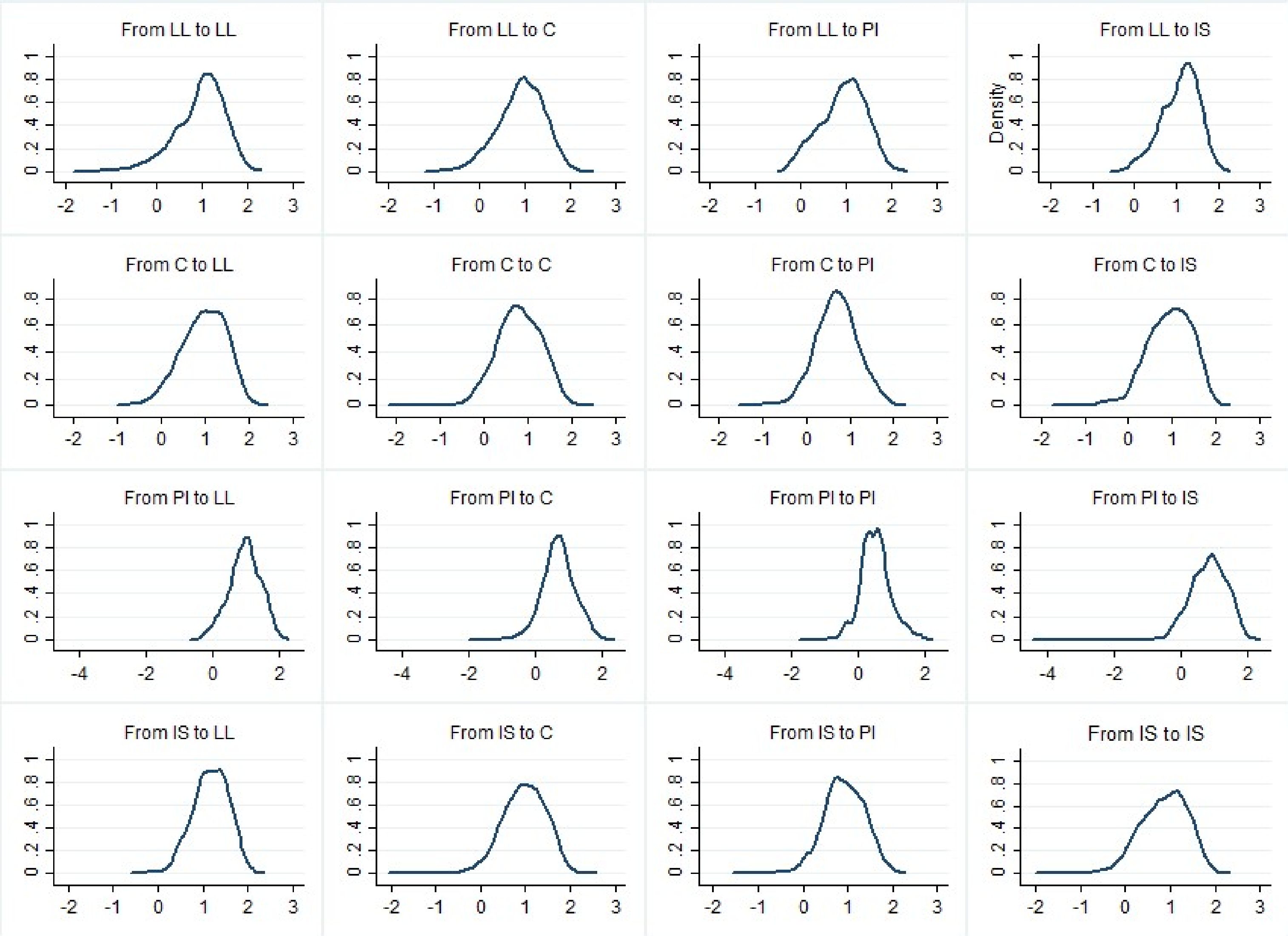
Table 4.4 Matrix of bilateral flows by geographical category: number and percent weight of positive flows, zero flows and asymmetric flows

i \ j	Landlocked	Coastal	Partial-insularity	Island-states
Landlocked	xij=0 & xji=0: **4068 (41%)** xij=0 & xji>0: **893 (9%)** xij>0 & xji=0: **853 (9%)** xij>0 & xji>0: **4106 (41%)**	xij=0 & xji=0: **8005 (29%)** xij=0 & xji>0: **2772 (10%)** xij>0 & xji=0: **2154 (8%)** xij>0 & xji>0: **14709 (53%)**	xij=0 & xji=0: **797 (18%)** xij=0 & xji>0: **426 (9%)** xij>0 & xji=0: **225 (5%)** xij>0 & xji>0: **3072 (68%)**	xij=0 & xji=0: **8459 (56%)** xij=0 & xji>0: **1360 (9%)** xij>0 & xji=0: **1200 (8%)** xij>0 & xji>0: **4141 (27%)**
Coastal	xij=0 & xji=0: **8497 (30%)** xij=0 & xji>0: **2546 (9%)** xij>0 & xji=0: **3105 (11%)** xij>0 & xji>0: **13892 (50%)**	xij=0 & xji=0: **14250 (19%)** xij=0 & xji>0: **5957 (8%)** xij>0 & xji=0: **5348 (7%)** xij>0 & xji>0: **49265 (66%)**	xij=0 & xji=0: **1277 (11%)** xij=0 & xji>0: **540 (5%)** xij>0 & xji=0: **467 (4%)** xij>0 & xji>0: **9336 (80%)**	xij=0 & xji=0: **17345 (43%)** xij=0 & xji>0: **2899 (7%)** xij>0 & xji=0: **4815 (12%)** xij>0 & xji>0: **15701(38%)**
Partial-insularity	xij=0 & xji=0: **851 (14%)** xij=0 & xji>0: **390 (6%)** xij>0 & xji=0: **521 (8%)** xij>0 & xji>0: **4598 (72%)**	xij=0 & xji=0: **1940 (11%)** xij=0 & xji>0: **733 (4%)** xij>0 & xji=0: **834 (5%)** xij>0 & xji>0: **14453 (80%)**	xij=0 & xji=0: **161 (6%)** xij=0 & xji>0: **64 (2%)** xij>0 & xji=0: **155 (6%)** xij>0 & xji>0: **2340 (86%)**	xij=0 & xji=0: **2284 (23%)** xij=0 & xji>0: **437 (5%)** xij>0 & xji=0: **1393 (14%)** xij>0 & xji>0: **5706 (58%)**
Island-states	xij=0 & xji=0: **9156 (50%)** xij=0 & xji>0: **1666 (9%)** xij>0 & xji=0: **2112 (12%)** xij>0 & xji>0: **5186 (29%)**	xij=0 & xji=0: **16688 (34%)** xij=0 & xji>0: **6497 (13%)** xij>0 & xji=0: **3659 (7%)** xij>0 & xji>0: **22876 (46%)**	xij=0 & xji=0: **1594 (20%)** xij=0 & xji>0: **924 (12%)** xij>0 & xji=0: **307 (4%)** xij>0 & xji>0: **5035 (64%)**	xij=0 & xji=0: **12336 (47%)** xij=0 & xji>0: **2136 (8%)** xij>0 & xji=0: **2483 (9%)** xij>0 & xji>0: **9565 (36%)**

Source: Author's elaborations from BACI-CEPII data, from the World Bank WDI database and from Pinna and Licio (2013) insularity data set

the sixteen possible combinations of bilateral trade between geographical categories, the number of flows and the corresponding percent weight (with respect to the total observations of the combination group) of bilateral positive trade flows, bilateral zero trade flows and asymmetric trade flows. The image coming from Table 4.4 highlights that flows involving landlocked countries or island-states are more prone of being zero. If one of the two trading partners is a country completely surrounded by the sea or completely surrounded by other countries, the case of absent trade is larger compared to trade between other geographical categories. When considering bilateral trade flows between two landlocked countries, the fraction of complete missing trade is 41 percent. This share is even higher for trade between two island-states (47 percent) and reaches half of all trade flows (50 and 56 percent) when trading partners are represented by a landlocked country and an island-state. These numbers suggest that landlocked countries and island-states face more difficulties in participating in global exchanges. In the perspective of the ʽunadjusted' measure of trade costs, Table 4.4 adds new knowledge, emphasising that high fractions of landlocked nations and island-states do not take part in international exchanges due higher trade costs they face. On this aim, Figure 4.1 reports the Kernel distribution densities of the log transformation of the ʽunadjusted' trade costs measure for all possible matches between geographical groups. The bandwidth has been set in order to obtain an optimally smoothed curve close to true density. Figure 4.1 shows that trade costs take negative values. As mentioned in Paragraph 4.1,

Figure 4.1 Kernel distribution of (unadjusted) trade costs (in logarithms) 1995-2014 by bilateral trade flows of geographical categories, bandwidth set optimally for each graph



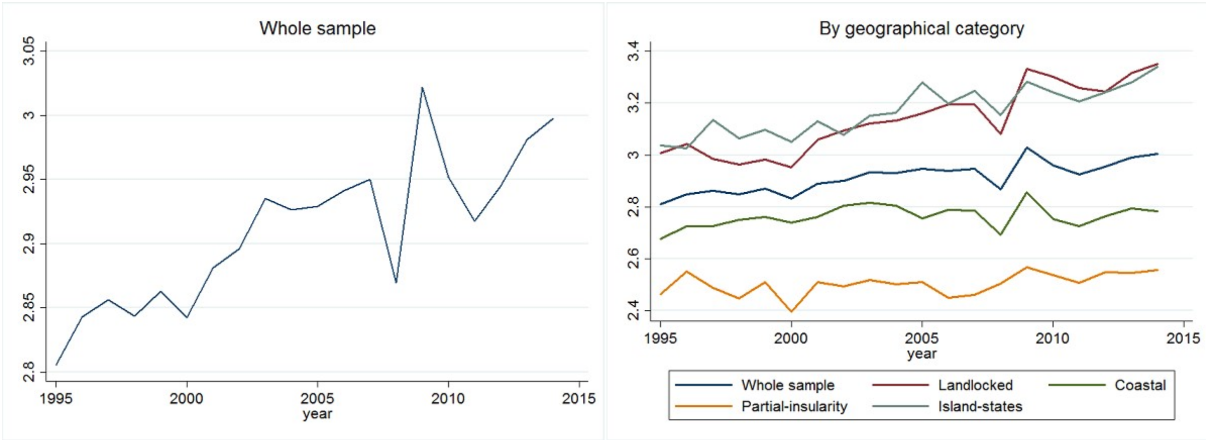
Source: Author's elaborations from BACI-CEPII data, from the World Bank WDI database and from Pinna and Licio (2013) insularity data set

trade costs not in logarithms may assume negative values when the internal trade for one or both partners is null or when internal trade is negative because of the high value of re-exports. None of the countries has internal trade equal to zero, but 11 out of 188 countries have negative internal trade.[15] However, the corresponding trade costs (not in logarithms) are negative only when one of the trading partner is Malaysia and Singapore. Moreover, the case of negative trade costs occurs also when none of the internal trades is negative, but the term $(x_{ii}*x_{jj})/(x_{ji}*x_{ji})$, greater than 1, at the power of $1/2(\sigma\text{-}1)$ is less than 1, producing in the end a negative trade cost (since the complete index requires to subtract 1). This case takes place only when the country pair is composed by the dyad Malaysia-Singapore. When considering the log transformation, as the Kernel densities do, the corresponding trade costs are negative for all those pairs for which trade costs not in logarithms are comprised between 0 and 1, and this represents 2 percent of total observations (country pairs). The picture that emerges from Figure 4.1 allows to appreciate that, when trade involves landlocked countries or island-states, trade costs tend to be higher, since a denser concentration of observations between 1 and 2 values. When considering trade between two partial-insularity countries, two coastal countries or a partial-insularity and coastal country, the density is higher between 0 and 1.

'*What has been the evolution of trade costs over time*' is one of the key issues that guided the opening of this thesis and that Figure 4.2 tries to answer looking at the world trade costs during last twenty years. Chapter 1 has underlined how during last thirty years international trade flows have increased extremely, and one of the sources explaining this incredible raise must be retrieved in the decline of trade costs. The increasing trend of the 'unadjusted' measure for both the whole sample and the single geographical groups in Figure 4.2 should be interpreted as a reduction in the share of absent bilateral flows due a reduction in trade costs. In 1995, the fraction of countries that do not trade at all was 75 percent higher compared to that in 2014.[16] Disregarding the trend and focusing on the values, it is possible to appreciate '*what are the differences between geographical groups*'. The right part of Figure 4.2 shows that landlocked countries and island-states experience the highest trade costs. Coastal countries are in the middle between the landlocked and island-states groups and the partial-insularity category. Unsurprisingly, due the very high fraction of rich economies, trade costs of partial-insularity countries are lower, compared to the other groups, and move differently from the other categories.

[15] Antigua and Barbuda, Aruba, Belize, Republic of Congo, Equatorial Guinea, Guyana, Liberia, Malaysia, Marshall Islands, Singapore, Tuvalu.
[16] In 1995 the number of zero bilateral flows was 8038, in 2014 4590.

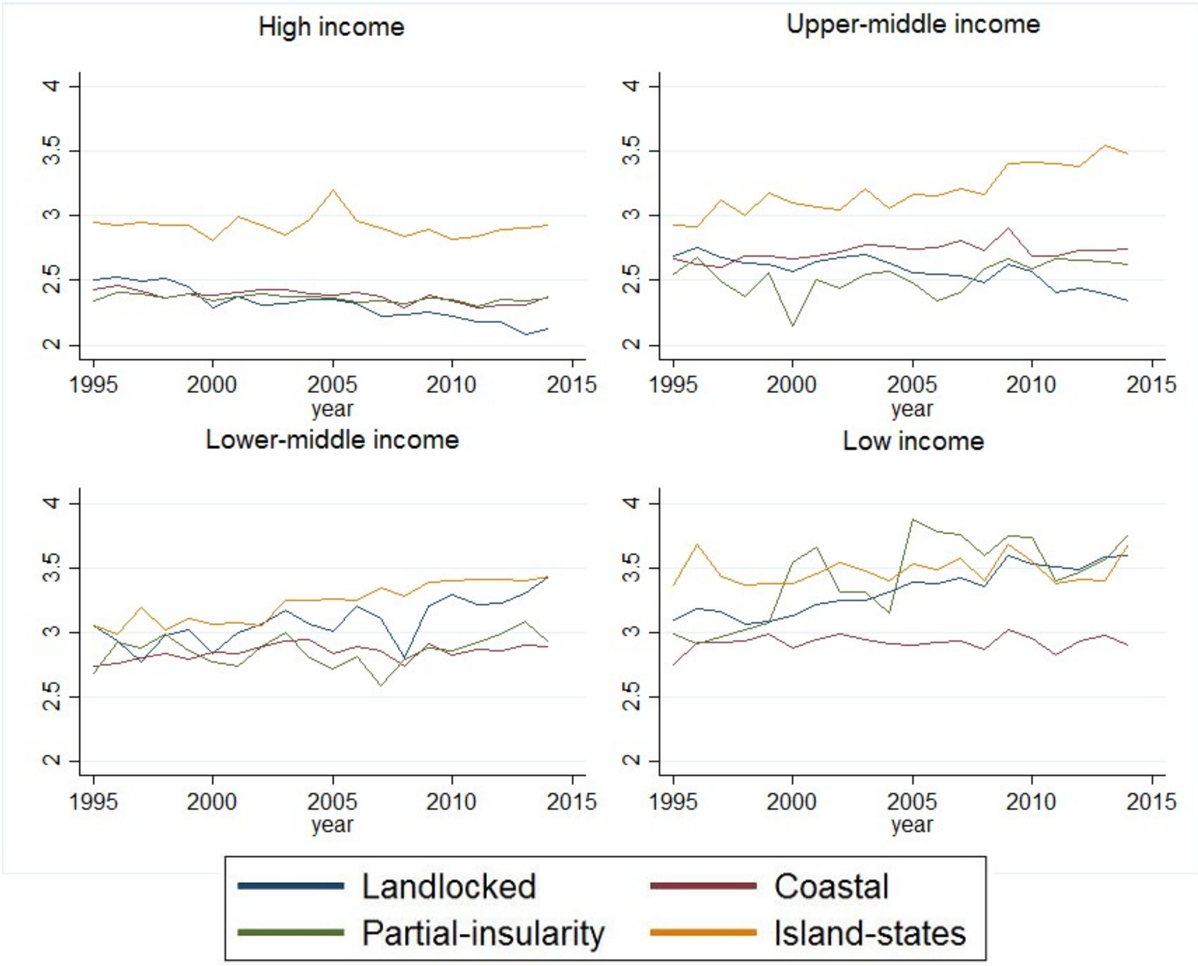
Figure 4.2 'Unadjusted' trade costs 1995-2014: whole sample and by geographical category



Source: Author's elaborations from BACI-CEPII data, from the World Bank WDI database and from Pinna and Licio (2013) insularity data set

From the descriptive so far performed, it clearly emerges how geography affects trade costs. Landlocked countries and island-states show the highest trade costs among the countries. Since the high fraction of poor countries within the landlocked group, it is hard to understand whether high trade costs originate from an unfortunate income condition, from a disadvantageous geographical state or from both. Island-states perform similarly to landlocked countries and for some years their trade costs are even higher. The case of island-states, which displays a more balanced representation in terms of income and where the fraction of rich countries weakly predominates, confirms that, for countries completely surrounded by the sea, the geography of being an island represents a disadvantage in terms of international exchanges, since it is associated to costly trade. For the other three geographical categories, a similar conclusion is not so obvious. On these bases, it is interesting to investigate '*what is the behaviour of trade costs when distinguishing countries by their WB income category*'. To better address this issue, Figure 4.3 connects both 'degree of insularity' and WB income category dimensions. When considering only rich countries, the most unfavourable geographical dimension is represented by being surrounded by the sea. The picture that emerges is that hostile geography, like being landlocked, is overcome when countries are rich, but the same countries are not able to overcome the adverse geography deriving from being an island. The case of upper-middle income countries shows a similar image. When looking at lower-middle income countries and at poor states, instead, the scene totally changes. Geographies of countries seem to be not so relevant anymore. At a first glance, not only countries seem to perform similarly, disregarding their advantageous or disadvantageous geography, but what is striking is that, within poor countries, the most

Figure 4.3 'Unadjusted' trade costs 1995-2014 by WB income category



Source: Author's elaborations from BACI-CEPII data, from World Bank data and from the World Bank WDI database

unfavourable condition is represented by having islands. It should be considered that only two countries (Equatorial Guinea and Guinea Bissau), within the partial-insularity group, are poor, hence this category is not adequately represented and does not present appropriate heterogeneity. Disregarding the partial-insularity category, once again island-states and landlocked countries have the most critical geographical conditions, but being an island is more disadvantageous than being landlocked. The descriptive emerged from this last figure adds an important evidence to the literature: whereas rich countries are able to overcome the adverse condition of having no access to sea, this is not the case when they are far from the mainland.

The impression coming from the descriptive obtained using the 'unadjusted' version of trade costs is that landlocked countries and island-states face more problems in taking part to global trade. These difficulties in some cases are so high that they impede trade completely. Even when they are able to participate in international exchanges, trade costs are higher compared to those of other geographies. On this view, an investigation on the main determinants of trade costs is appropriate, in order to understand what makes trade more costly in those countries characterised by an unfavourable geography.

4.3.2. The potential main determinants of trade costs

The choice of the possible main factors influencing trade costs has considered, on the one hand, the previous findings coming from the trade costs literature (Arvis et al., 2013a, 2013b; Duval and Utoktham, 2011b) and, on the other hand, the typical factors that normally enhance competitiveness and ensure future growth and further development.[17] On this regard, trade costs main determinants can be attributed to four main spheres: i) geography; ii) logistics, connectivity and competitiveness; iii) infrastructure; iv) documentary, border and transport compliance.

Within the geographical dimension, three basic measures have been selected to capture the countries' geography: ruggedness, average temperature and distance from the Equator.

The terrain ruggedness index, originally developed by Riley, DeGloria and Elliot (1999) to quantify topographic heterogeneity in wildlife habitats creating hiding places and outlook posts, comes from Nunn and Puga (2012) data set and quantifies small-scale terrain irregularities. It is measured in hundreds of metres of elevation difference for grid points 30arc-seconds (926 metres on a meridian) apart and computed as the square root of the sum of the squared differences in elevation between a central point and eight adjacent points. The idea behind the use of this index is that it is able to capture the economic disadvantage of countries marked by a higher ruggedness of their land. The information is available for 234 countries in the world.

The second geographical variable is represented by the average temperature, computed using the data on the maximum and minimum annual temperature by country, provided by the World Bank (Climate Change Knowledge Portal) for 201 economies in the world, in degrees Celsius. The insight of using the average temperature comes from the literature: temperate areas have a positive effect on institutions and on health conditions, fostering growth and development.

[17] To understand why trade costs are high or simply why there are trade costs, it is fundamental to understand first what makes trade costly and where these costs arise. This issue has been excellently developed by Moïsé and Le Bris (2013), which review and summarise all OECD research and findings on cost factors. They place along the entire trade chain (behind the border, crossing the border or beyond the border) all those factors that make trade costly.

Lastly, since the evidence coming from economic research, according to which smaller latitude is correlated with bad institutions, the distance from Equator is added as third geographical feature able to explain trade costs pattern. Data are from La Porta et al. (1997) and include the distance latitude of the country's capital from the Equator for 207 countries.

The sphere linked to logistics, connectivity and competitiveness includes three important indexes with a high potential explanatory power. The first one is the Logistic Performance Index (LPI) provided by the World Bank (WDI, Logistics Performance Index surveys) for 192 countries[18] and for four years (2007, 2010, 2012, 2014). The LPI is described as the "*perceptions of a country's logistics based on efficiency of customs clearance process, quality of trade- and transport-related infrastructure, ease of arranging competitively priced shipments, quality of logistics services, ability to track and trace consignments, and frequency with which shipments reach the consignee within the scheduled time. The index ranges from 1 to 5, with a higher score representing better performance*".[19] The respondents have been asked to evaluate the eight most important export and import markets of the respondent's country on six core dimensions. The first two columns of Table 4.5 rank countries according to their LPI, showing the first five, with the highest LPI, and the last five states, with the lowest LPI, in 2007 (initial year for which data are available) and 2014 (last year existing). As expected, Western Europe countries occupy the first places, but it is Singapore that performs better in 2007. In 2014, the picture does not change: again European countries at the top of the list, but with now Germany in first place and Singapore in the fifth one. The curious fact lies in the worst countries of 2007: at the top of the list, four out of five nations are Asian countries. In 2014, a perfectly reverse image emerges: four out of five countries are from the Sub-Saharan Africa, with Afghanistan in third worst position.

The second index included in the category, that this chapter has named 'logistics, connectivity and competitiveness', is the Linear Shipping Connectivity Index (LSCI) provided by the United Nations Conference on Trade and Development (UNCTAD) for the period 2004-2016 and for 159 countries in the world. The LSCI captures how well countries are connected to global shipping networks, and it is based on five components of the maritime transport sector: number of ships, container-carrying capacity, maximum vessel size, number of services, and number of companies that arrange container ships in a country's ports. The index is as an average of all five components and yields a value of 100 for the country with the highest

---

[18] The data set is composed by 248countries, but the information is available for 182 countries in 2007, 187 in 2010, 186 in 2012 and 192 in 2014.

[19] World Bank, WDI (http://data.worldbank.org/indicator/LP.LPI.OVRL.XQ). For more information see Arvis et al. (2014) "Connecting to Compete 2014: Trade Logistics in the Global Economy", World Bank.

average index in 2004 (UNCTAD, 2016). The underlying data come from Containerisation International Online (World Bank, WDI 2016). The insight of using the LSCI as determinants of trade costs comes from the triple circular relationship between trade, transport costs and transport services. Better transport services determine more trade, more money to finance new infrastructure and, in turn, better transport services. Lower transport costs determine more trade, higher economies of scale and, in turn, lower transport costs. More trade determines more shipping supply, fostered competition, lower shipping costs and, in turn, more trade (UNCTAD, 2009).[20] The third and fourth columns of Table 4.5 show the first five and the last five countries according to their LSCI. In order to be comparable with LPI ranks, and since next inference analysis will be based mainly on the years for which the information about the LPI is available, data have been selected for 2007 and 2014. From the first five positions, it is immediately apparent that, both in 2007 and 2014, Asian countries have the best maritime shipping connectivity, with China leading the rank in both years. The five countries having the lowest indexes in 2014 are completely small and remote islands or archipelagos. In 2007, the rank of the worst LSCI also includes an African country (Eritrea) and Greenland.

Table 4.5 Top five and worst five countries by Logistic Performance Index, Linear Shipping Connectivity Index and Global Competitiveness Index in 2007 and 2014

Logistic Performance Index				Linear Shipping Connectivity Index				Global Competitiveness Index			
2007		2014		2007		2014		2007		2014	
Rank	Country	Rank	Country	Rank	Country	Rank	Country	Rank	Country	Rank	Country
1	Singapore	1	Germany	1	China	1	China	1	United States	1	Switzerland
2	Netherlands	2	Netherlands	2	Hong Kong	2	Hong Kong	2	Switzerland	2	Singapore
3	Germany	3	Belgium	3	Germany	3	Singapore	3	Denmark	3	United States
4	Sweden	4	United Kingdom	4	Singapore	4	Korea, Rep.	4	Sweden	4	Finland
5	Austria	5	Singapore	5	Netherlands	5	Malaysia	5	Germany	5	Germany
182	Afghanistan	192	Somalia	159	Eritrea	159	Cayman Islands	127	Chad	143	Guinea
181	Timor-Leste	191	Congo, D.R.	158	Bermuda	158	Palau	126	Burundi	142	Chad
180	Rwanda	190	Afghanistan	157	Sao Tome & P.	157	Micronesia	125	Zimbabwe	141	Yemen
179	Myanmar	189	Congo, Rep.	156	Cayman Islands	156	Bermuda	124	Mozambique	140	Mauritania
178	Tajikistan	188	Eritrea	155	Greenland	155	Dominica	123	East timor	139	Angola

Source: Author's elaborations from the World Bank WDI database, from UNCTAD data, from the Global Competitiveness Index Historical Dataset 2005-2015 World Economic Forum

The third measure refers to the Global Competitiveness Index (GCI) developed by Xavier Sala-i-Martín from the Global Competitiveness Index Historical Dataset 2005-2015 - World Economic Forum. Data are structured in ten editions (from edition 2006-2007 to edition 2015-

[20] For more information see “Review of Maritime Transport 2009” (UNCTAD, 2009) and “Review of Maritime Transport 2016 - The long-term growth prospects for seaborne trade and maritime businesses” (UNCTAD, 2016).

2016)[21] and available for 147 countries.[22] The GCI captures country performance since 2004, combining 114 indicators strictly connected with productivity. These indicators are grouped into 12 pillars: institutions, infrastructure, macroeconomic environment, health and primary education, higher education and training, goods market efficiency, labour market efficiency, financial market development, technological readiness, market size, business sophistication, and innovation. These 12 pillars are in turn structured in three sub-indexes, in line with three main stages of development: basic requirements, efficiency enhancers, and innovation and sophistication factors (World Economic Forum, 2015).[23] The last two columns of Table 4.5 rank countries according to their GCI. Once again, Western European countries prevail in the top part of the list. Last positions are, instead, almost totally occupied by African countries. The U.S. are the most competitive country in the world in 2007, according to the GCI; Switzerland predominates in 2014. Chad and Guinea are the worst countries in competitiveness terms in 2007 and 2014, respectively.

Figure 4.4 Logistic Performance Index, Linear Shipping Connectivity Index and Global Competitiveness Index in 2007 and 2014: whole sample (WS) and by geographical category (LL=Landlocked; C=Coastal; PI= Partial-insularity; IS=Island-states)

Source: Author's elaboration from the World Bank WDI database, UNCTAD data, the Global Competitiveness Index Historical Dataset 2005-2015 World Economic Forum and from the Pinna and Licio (2013) insularity data set

21 To allow descriptive and empirical analysis, editions have been simply converted into years. Edition 2006-2007 has been referred to year 2006, edition 2007-2008 to year 2007, and so on.

22 The data set is composed by 150 countries, with some missings. For the edition 2006-2007 data are available for 120 countries, for the edition 2007-2008 127 countries, for the edition 2008-2009 and the edition 2009-2010 131 countries, for the edition 2010-2011 138 countries, for the edition 2011-2012 141 countries, for the edition 2012-2013 143 countries, for the edition 2013-2014 147 countries, for the edition 2014-2015 143 countries, and for the edition 2015-2016 139 countries.

23 For more information see “The Global Competitiveness Report 2015–2016” (World Economic Forum, 2015).

Figure 4.4 shows the three indexes explained above by geographical category in 2007 and 2014. Apart from the significant increase of the LSCI from 2007 to 2014, there are not noteworthy differences between the two years. What immediately emerges is the best performance of the partial-insularity group for each of the three indexes, but also the good logistics abilities of the island-states group. In fact, when considering the LPI, the island-states have on average better-quality logistics aptitudes than coastal countries. For the LSCI, the picture is completely the opposite: coastal countries perform better than island-states and this difference enlarges in 2014. When looking at the GCI, there are not remarkable diversities between coastal and island countries, although, once again, the group of the island-states has higher average values. The landlocked countries have lower values for both LPI and GCI; since they are landlocked, the LSCI cannot be obviously measured.

Table 4.6 Top five and worst five countries by quality of roads, quality of port infrastructure and quality of air infrastructure in 2007 and 2014

Quality of roads				Quality of ports				Quality of air infrastructure			
2007		2014		2007		2014		2007		2014	
Rank	Country	Rank	Country	Rank	Country	Rank	Country	Rank	Country	Rank	Country
1	France	1	United Arab E.	1	Singapore	1	Netherlands	1	Singapore	1	Singapore
2	Singapore	2	Portugal	2	Netherlands	2	Singapore	2	Hong Kong	2	United Arab E.
3	Switzerland	3	Austria	3	Hong Kong	3	United Arab E.	3	Germany	3	Hong Kong
4	Germany	4	France	4	Germany	4	Hong Kong	4	United Arab E.	4	Netherlands
5	Denmark	5	Netherlands	5	Denmark	5	Finland	5	France	5	Finland
127	Chad	143	East Timor	127	Tajikistan	143	Kyrgyzstan	127	Lesotho	143	Lesotho
126	Mongolia	142	Guinea	126	Kyrgyzstan	142	Mongolia	126	East Timor	142	East Timor
125	East Timor	141	Libya	125	Bosnia & Herzeg.	141	Chad	125	Paraguay	141	Chad
124	Bosnia & Herzeg.	140	Mozambique	124	Albania	140	Bolivia	124	Cameroon	140	Yemen
123	Paraguay	139	Moldova	123	East Timor	139	Tajikistan	123	Bosnia & Herzeg.	139	Mauritania

Source: Author's elaborations from the Global Competitiveness Index Historical Dataset 2005-2015 World Economic Forum

The third set of potential trade costs determinants refers to the infrastructure area. The Global Competitiveness Index Historical Dataset 2005-2015 of the World Economic Forum includes, among others, four measures that refer to the quality of infrastructure. These measures come from the Executive Opinion Survey, which determines those aspects that require a more quality evaluation or for which internationally comparable statistical data are not available, and captures the opinions of business leaders around the world on a wide range of topics for which data are scarce or do not exist on a global scale (World Economic Forum, 2015).[24] For the purposes of this chapter, four indicators have been chosen: quality of roads, quality of port

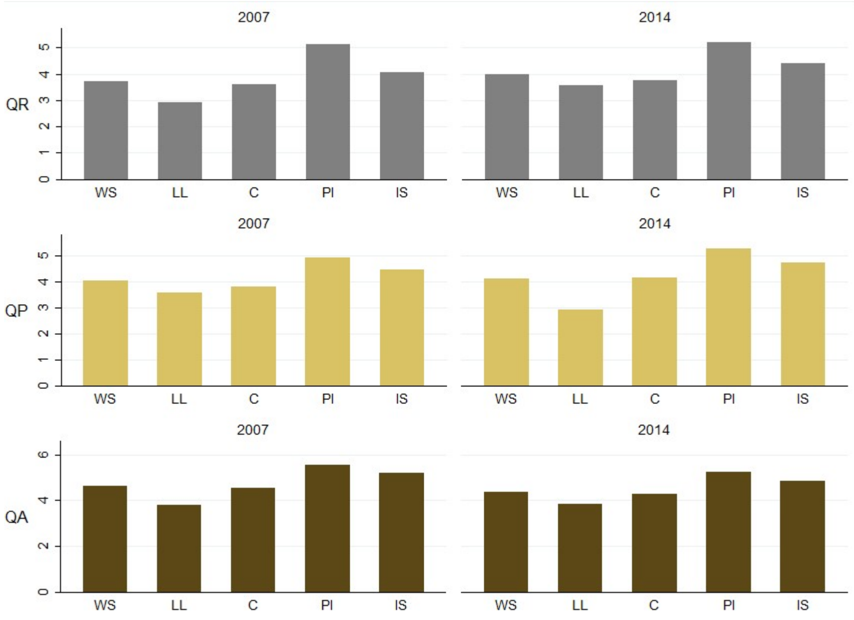
[24] For more information about the Survey see "The Global Competitiveness Report 2015–2016" (World Economic Forum, 2015).

infrastructure, quality of air infrastructure and quality of overall infrastructure. The measures range from 1 (extremely underdeveloped, among the worst in the world) to 7 (extensive and efficient, among the best in the world). Respondents have been asked to answer the following questions: "*In your country, how do you assess the quality of the roads?*"; "*In your country, how do you assess the quality of seaports (for landlocked countries, assess access to seaports)?*"; "*In your country, how do you assess the quality of air transport?*"; "*How do you assess the general state of infrastructure (e.g., transport, communications, and energy) in your country?*". Table 4.6 ranks countries according to the quality of their roads, ports and air infrastructures. What immediately comes into sight is how for both 2007 and 2014 a small group of European and Asian countries predominate the top five of all three rankings. Singapore appears in the top five ranks five times out of six: it comes in first place in 2007 and in second place in 2014 for the highest quality of ports, and in first place in both years for the best quality of its air infrastructure; moreover, in 2007, it is in second place for the quality of roads. The United Arab Emirates compare four times: its roads are the best in 2014, but the quality of its ports and air infrastructure is also very highly assessed. The Netherlands, Hong Kong, France, Germany, Denmark and Finland have more than one infrastructure that is well considered by respondents. On the contrary, among the worst there is a very high variability: Sub-Saharan Africa, South and East Asia, Middle-East and North Africa countries, but also European and Latin America nations compare in the ranks. Bosnia and Herzegovina is the European state with the lowest quality of land, sea and air infrastructure. Albania faces some insufficiency in the port infrastructure, Moldova in the roads. Within the group of South America countries, Paraguay and Bolivia are marked by some deficiency. Figure 4.5 shows the three indexes explained above by geographical category in 2007 and 2014: there are not noteworthy differences in the quality of infrastructure between the two years.

In the fourth set of determinants of trade costs, there are those measures which have to do with the documentary, border and transport compliance. Data are from the Doing Business database of the World Bank, which provides information for 186 countries[25] in the world and for a quite long time-span (2004-2016). Seven measures related to trading across borders have been selected. Data on trading across borders are assembled through a questionnaire administered to local freight forwarders, customs brokers, port authorities and traders (World Bank, 2016). The first measure refers to the trading across border score (TAB), which captures "*the time and cost (excluding tariffs) associated with three sets of procedures*

[25] The data set is intended for 186 countries, but not for all years the information is available for all 186 countries.

Figure 4.5 Quality of roads (QR), quality of port infrastructure (QP) and quality of air infrastructure (QA) in 2007 and 2014: whole sample (WS) and by geographical category (LL=Landlocked; C=Coastal; PI= Partial-insularity; IS=Island-states)



Source: Author's elaboration from the Global Competitiveness Index Historical Dataset 2005-2015 World Economic Forum and from the Pinna and Licio (2013) insularity data set

(documentary compliance, border compliance and domestic transport) within the overall process of exporting or importing a shipment of goods". The indicator is a distance to frontier score, measuring the distance of each economy to the frontier, which represents the best performance observed in each of the indicators across all economies in the Doing Business sample since 2005. "*An economy's distance to frontier is reflected on a scale from 0 to 100, where 0 represents the lowest performance and 100 represents the frontier*".[26] The other six measures selected refer to all those six spheres that are considered when computing the trading across border score: documents to export, time to export, cost to export, documents to import, time to import and cost to import. Among the measures reflecting the documentary compliance, the number of documents needed to export and the number of documents needed to import have been selected. Within the group of measures related to the time to trade, the number of days to export and the number of days to import have been chosen. Finally, among the indexes that capture the cost to trade, the US$ per exporting container and the US$ per importing container have been selected.[27] Table 4.7 summarises all seven measures for the whole sample and for each geographical category in the two selected years 2007 and 2014.

[26] For more information see "Doing Business 2016: Measuring Regulatory Quality and Efficiency" (World Bank, 2016).
[27] For the descriptive analysis the US$ per exporting container and the US$ per importing container, refer to deflated values. For the inference analysis, both variables refer to current (not deflated) values.

From a general perspective, it comes out that the comprehensive trading across borders measure has improved from 2007 to 2014. This means that the distance to frontier has reduced and that all those compliances related to crossing the boundary have become simpler and faster. In fact, comparing the six selected measures which contribute, among others, to the trade across border score, it is immediately clear how all single measures have experienced a decrease during the seven years between 2007 and 2014. Moreover, it also comes into sight how the number of documents and of days to export and the cost to export a container are always lower compared to the documents, the days and the costs to import. From the geography point of view, once more, the group of partial-insularity countries perform better, followed by the group of the island-states, by coastal countries and, lastly, by landlocked countries.

Table 4.7 Trading Across Borders measures in 2007 and 2014: whole sample and by geographical category

		Whole Sample	Landlocked	Coastal	Partial-insularity	Island-states
Trading Across Borders	2007	58.27	32.28	59.35	78.40	69.26
	2014	66.48	40.80	68.16	81.28	75.96
N° documents to export	2007	6.62	8.65	6.64	4.75	5.70
	2014	6.22	7.90	6.31	4.56	5.48
N° documents to import	2007	7.96	10.84	7.89	5.19	6.86
	2014	7.34	9.61	7.35	4.88	6.61
N° days to export	2007	26.88	45.06	24.71	15.31	20.89
	2014	21.87	38.77	19.88	13.63	16.82
N° days to import	2007	31.10	52.45	30.07	16.44	21.54
	2014	24.40	43.48	23.12	14.13	17.18
Cost to export (deflated US$ per container)	2007	2298.52	4153.37	2184.10	1248.43	1421.21
	2014	1574.95	3231.88	1381.16	1111.53	950.45
Cost to import (deflated US$ per container)	2007	2754.09	5147.91	2559.30	1313.23	1750.59
	2014	1889.21	3897.05	1655.07	1170.98	1188.06

Source: Author's elaborations from the World Bank Doing Business Database and from the Pinna and Licio (2013) insularity data set

From this double descriptive analysis, trade costs on the right-hand side and all regressors on the left-hand side, there is a very interesting fact that emerges. Although island-states have trade costs comparable to those of landlocked countries and much higher than those of coastal countries, when looking at the single measures of logistics, competitiveness and performance, island-states perform highly better than coastal countries, suggesting that there is something more related to the their geography which affects trade costs. In fact, when considering the sea shipping connectivity, some difficulties emerge.

4.4 The empirical model

As stated above and as also stressed by Arvis et al. (2013a), the main challenge of the analysis, aimed at assessing the main determinants of trade costs, is represented by the data constraints. Additionally, the complexity of matching different data sets with different countries, different time spans and different missing countries in different years further complicates the investigation, yielding additional drawbacks for the inferential analysis.

An extra limitation is represented by the time span. Unfortunately, the long 20 years period for which the information about trade costs is available cannot be used in the empirical analysis, since the independent variables refer mainly to very recent years. Table 4.8 summarises the data and the sources used to construct the variables included in the analysis. As it is possible to notice from Table 4.8, the main drawback is represented by the Logistic Performance Index (LPI), that is available for just four years. Since previous findings (Arvis et al., 2013a) show the high importance of this regressor in the ability to explain trade costs, the simple choice to sacrifice the LPI cannot be accomplished and the inference analysis will focus on the years for which the variable is available. Another difficulty is represented by the Linear Shipping Connectivity Index (LSCI): this regressor is absent for those countries that don't have access to the sea; hence, when including this variable in the model, the analysis is forced to renounce to the landlocked countries, which, it must be reminded, are mainly poor countries. In this view, when the model includes the LSCI, it explains trade costs only for those countries that are not landlocked.

The descriptive so far performed has shown how 31 percent of total flows are represented by cases of complete missing trade (i.e. both flows are zero), 17 percent are cases of asymmetric trade flows (i.e. one flow is zero and one is positive), and 52 percent are cases of positive trade (i.e. both flows are higher than zero). In this light, the empirical analysis aims, on the one hand, to explore what affects the participation of countries to international exchanges and, on the other hand, to examine what makes trade costly. The ‛unadjusted’ version of trade costs is the most suitable for this kind of investigation, since trade costs are missing when trade flows are zero and should be interpreted as non-participation to international exchanges due a very high value of the cost to trade. However, the ‛unadjusted’ measure does not distinguish between cases where trade is missing in both directions and cases where just one country does not export to the trading partner.

Founding the indirect measure of trade costs on the gravity model entails the need to face the econometric weaknesses deriving from the cases of zero trade flows. The main

Table 4.8 Variables and data sources

Dependent variable	Definition	Years	Source	Countries in the data set	Countries info available	Other info
Trade costs	Geometric average of international trade costs between countries i and its partners js relative to domestic trade costs within each country	1995-2014	BACI-CEPII and WDI	188	188	Not 188 countries for all years
Distance	Weighted distance in km between country i and j (weight: population)	time invariant	BACI-CEPII	pairs	188	
Ruggedness	Terrain Ruggedness Index, 100 m	time invariant	Nunn and Puga (2012	234	234	
Average temperature	Average temperature in degrees Celsius.	time invariant	World Bank, Climate Change Knowledge Portal	211	201	
Distance from equator	Distance latitude of the country's capital to from the Equator	time invariant	La Porta et al. (1997)	208	207	
Landlocked	1 if the country has not access to the sea	time invariant	Pinna and Licio (2013)	232	232	37 landlocked
Coastal	1 if the country has access to the sea and the insular territory is equal or less than 2% of the total territory	time invariant	Pinna and Licio (2013)	232	232	94 coastal
Partial insularity	1 if the country has access to the sea and the insular territory is more than 2% of to the total territory	time invariant	Pinna and Licio (2013)	232	232	17 partial insularity
Dummy island-state	1 if the country is an island	time invariant	Pinna and Licio (2013)	232	232	84 island-states
Linear Shipping Connectivity Index	How well countries are connected to global shipping networks	2004-2016	UNCTAD	159	159	
Logistic Performance Index	Efficiency of customs clearance process, quality of trade and transport-related infrastructure, ease of arranging competitively priced shipments, quality of logistics services, ability to track and trace consignments, frequency with which shipments reach the consignee within the scheduled time.	2007, 2010, 2012, 2014	World Bank, WDI	248	192	Not 192 countries for all years
Global Competitiveness Index	Combines 114 indicators that capture institutions, infrastructure, environment, health, education, goods market efficiency, labor market efficiency, financial market development, technological readiness, market size, business sophistication, and innovation	2006-2015	The Global Competitiveness Index Historical Dataset World Economic Forum	150	150	Not 150 countries for all years
Quality of overall infrastructure	Survey: 1 = extremely underdeveloped among the worst in the world; 7 = extensive and efficient among the best in the world	2006-2015	The Global Competitiveness Index Historical Dataset World Economic Forum	150	150	Not 150 countries for all years
Quality of roads	Survey: 1 = extremely underdeveloped among the worst in the world; 7 = extensive and efficient among the best in the world	2006-2015	The Global Competitiveness Index Historical Dataset World Economic Forum	150	150	Not 150 countries for all years
Quality of ports	Survey: 1 = extremely underdeveloped among the worst in the world; 7 = extensive and efficient among the best in the world	2006-2015	The Global Competitiveness Index Historical Dataset World Economic Forum	150	150	Not 150 countries for all years
Quality of air infrastructure	Survey: 1 = extremely underdeveloped among the worst in the world; 7 = extensive and efficient among the best in the world	2006-2015	The Global Competitiveness Index Historical Dataset World Economic Forum	150	150	Not 150 countries for all years
Trading across borders	Time and cost (excluding tariffs) associated with documentary compliance, border compliance and domestic transport within the overall process of exporting or importing a shipment of goods. Distance to frontier.	2004-2016	Doing Business database, World Bank	186	186	Not 186 countries for all years
Documents to export	Number of documents	2004-2016	Doing Business database, World Bank	186	186	Not 186 countries for all years
Time to export	Number of days	2004-2016	Doing Business database, World Bank	186	186	Not 186 countries for all years
Cost to export	US$ per container	2004-2016	Doing Business database, World Bank	186	186	Not 186 countries for all years
Documents to import	Number of documents	2004-2016	Doing Business database, World Bank	186	186	Not 186 countries for all years
Time to import	Number of days	2004-2016	Doing Business database, World Bank	186	186	Not 186 countries for all years
Cost to import	US$ per container	2004-2016	Doing Business database, World Bank	186	186	Not 186 countries for all years
WB income category	4 groups according to gross national income per capita of the previous year	1995-2014	World Bank	211	211	Not 211 countries for all years
Common language	1 if the country pair has a common official primary language	time invariant	CEPII	pairs	188	
Common legal origin	1 if the country pair has a common legal origin	time invariant	CEPII	pairs	184	

Source: Author's elaborations

problem when dealing with zero trade flows is represented by the bias of estimates which mainly originates from the sample selection. And this can be better understood when explaining the zero trade issue and the connected sample selection bias looking at two facts: heteroskedastic residuals and prevalence of zeros in the dependent variable. Santos Silva and Tenreyro (2006) clearly address these two issues. First, heteroskedastic residuals are important in the gravity framework not only for the link with zeros, but because of the common practice of log-linearising the gravity equation.[28] When considering the zeros issue, instead, the connection between heteroskedasticity and zero trade flows arises via the sample selection. Hurd (1979) points out that heteroskedasticity may rise biased estimates in truncated samples, i.e. where zero trade flows are excluded. In censored sample, i.e. where zero trade flows are kept, instead, Arabmazar and Schmidt (1981) explain this is less the case. The second point has to do with the commonness of zeros in the dependent variable. Santos Silva and Tenreyro (2006) highlight how, in the gravity framework, the prevalence of zero trade flows and the combined fact of usually expressing the dependent variable in logarithms leads to undefined values and produces the risk of having biased estimates, since this force of dealing with a truncated sample. In fact, OLS automatically drops these observations, since the logarithm of zero is undefined.

The literature on the gravity model has dealt with zero trade flows, adopting four different estimation solutions. McCallum (1995) and Frankel (1997) drop observations with zero yielding a truncated sample. The second method replaces zero trade flows with small positive constants, according to Linnemann (1966), McCallum (1995) and Raballand (2003). Rose (2000), Soloaga and Winters (2001), Anderson and Marcouiller (2002) and Baldwin and DiNino (2005) exploit a third solution using a Tobit (1958) model, which censors part of the observations to zero, or rounds up zero trade flows below a given positive value. The fourth method, exploited among others by Emlinger et al. (2008), Disdier and Marette (2010) and Jayasinghe et al. (2010), refers to the Heckman (1979) model, which corrects the sample selection bias exploiting a two-stage econometric approach.

---

[28] The practice of log-linearising the gravity equation and of estimating it using OLS, yields biased estimates. As highlighted by Santos Silva and Tenreyro (2006), this comes from the Jensen's Inequality, according to which the expected value of a logarithm is different from the logarithm of an expected value: $E(\ln(x)) \neq \ln(E(x))$. With the logarithmic transformation, the error term is not u_i anymore, it is $\ln(u_i)$, producing estimates of $E(\ln(u_i))$ and not estimates of $\ln(E(u_i))$. This fact makes OLS estimates unbiased. The OLS estimator is able to produce unbiased parameters only if the error term of the log-linearised specification is homoskedastic, i.e. the variance of the error term is independent of the covariates: $\ln(E(u_i|x_i)=0$. But, in log-linearised gravity equation the property of the error term changes. It becomes heteroskedastic, i.e. the variance of the error term is a function of regressors: $E(\ln(u_i|x_i)\neq 0$. In other terms, since in log-linearised specification the error term is $\ln(u_i)$ and not u_i, since $E(\ln(u_i))$ is different from $\ln(E(u_i))$ and since $E(\ln(u_i|x_i)\neq 0$ is different from $\ln(E(u_i|x_i)=0$, then OLS estimator cannot yield unbiased estimates.

As noticed by Kareem (2013), the Tobit (1958), the Heckman (1979) and the Helpman, Melitz and Rubinstein (2008) models are, in the gravity literature, the most appropriate approaches to deal with the estimation drawbacks that originate from zero trade flows and logarithm transformation. However, De Benedictis and Taglioni (2011) advise that choosing the best estimation technique is not simple and it may vary according to researchers' needs. The way according to which data are structured and organised, the kind of assumptions that the analysis imposes and the level of overdispersion determine the selection of the most suitable method.

In the trade costs framework, and in particular in the indirect trade costs framework, due the close and deep connection with the gravity equation, having zero trade flows yields the same weaknesses presented above for the gravity model. On this purpose, in order to avoid biased estimates, in order to take into account the participation to trade and to explore both extensive and intensive margins,[29] the ideal approach exploited is the one proposed by Heckman in 1979 and that refers to the Heckman model.[30]

The Heckman (1979) model is a two-steps estimation procedure that enables to control for sample selection bias. Moreover, it allows to control for the omitted variable bias: the probability of being selected for the estimation sample (due positive trade flows) can be interpreted as omitted variable, since the fact of having positive trade flows is due to lower trade costs which affect the extent of trade.

In the gravity framework, the typical first step of the Heckman (1979) model consists in estimating the probability of exporting (dependent variable) using a Probit model. This first equation is also named *selection equation*, since it investigates the relationship between the probability of positive trade and a series of covariates able to explain it. This first stage allows to explore the extensive margin of trade (i.e. the number of trade relationships between countries). The second step is more focused on the intensive margin (i.e. the number of exporting firms) and consists in estimating an *outcome equation*, i.e. a standard gravity equation in logarithms where only positive trade (dependent variable) is included in the estimation.

The selection equation comprises what is called *latent variable* or *excluded variable*. It consists in a single variable or a set of explanatory variables that are included in the selection but not in the outcome equation: they affect the probability to trade (dependent variable of the

[29] Xiong and Beghin (2011) point out how the Heckman model allows to explore both extensive and intensive margins of trade.

[30] Robustness analysis solves the problem of biased estimates using the 'adjusted' version of trade costs, that, with the substitution of zero trade flows with one, yields a censored sample.

selection equation) but do not affect the volume of trade (dependent variable of the outcome equation).

When the set of covariates in both selection and outcome model is the same, then the Heckman model is just-identified. However, an over-identified model (i.e. the number of explanatory variables in the selection regression are greater than those included in the outcome regression) is more appropriate. In fact, the inclusion in the selection model of one or more exclusion restrictions enables to achieve a better identification. These excluded variables are used to compute the inverse Mills ratio,[31] which represents the probability of the latent variables excluded from the outcome regression. The inverse Mills ratio is then included in the outcome regression as an additional explanatory variable.

The Heckman (1979) model relies on two different estimators: the full Maximum Likelihood estimator[32] and the Two-Step consistent estimator. On the one hand, as found by Puhani (2000), the Maximum Likelihood estimator yields better results than the Two-Step estimator. Leung and Yu (1996) argue that a limited number of exclusion restrictions, low variability among regressors, a high variance of the error term and a strong censored sample limit the performance of the Two-Step estimator. Nevertheless, it is more appropriate with large data sets.[33]

For the purpose of this chapter, the Heckman model, together with the 'unadjusted' version of trade costs, considers, in the first stage, missing trade costs (i.e. when one or both bilateral flows are zero) as zeros, interpreting them as non-participation to trade, and positive trade costs as ones, considering them as being engaged in trade. t_{ij} is the non-observed dependent variable of the selection model that is observed if trade costs are missing or positive. In the second stage of the procedure, only positive[34] trade costs are explored, and the regression analysis allows to investigate the main and potential determinants of trade costs. The natural logarithm of τ_{ij} is the dependent variable of the outcome model and represents the natural logarithm of positive trade costs.

[31] The inverse Mill's ratio is the ratio of the probability density function over the cumulative distribution function of a distribution. It is determined by the arc hyperbolic tangent of ρ and by the logarithm of standard errors of the selection equation.

[32] Maximum Likelihood estimator requires to assume that $u_i \sim N(0,1)$, $\epsilon_i \sim N(0,\sigma^2)$, and that $corr(u_i, \epsilon_i) = \rho$.

[33] StataCorp (2013).

[34] Although the ordinary case of positive trade costs, two more unusual cases may occur: trade costs equal to zero or negative trade costs. The situation of zero trade costs arises when in both countries the internal trade is equal to the exports to the trading partner. The case of negative trade costs occurs when the internal trade for one or both partners is null (i.e. total exports equal to GDP) or, as highlighted by De Benedictis and Pinna (2015), when internal trade is negative because of the high value of re-exports. On this view, it can be argued that trade costs should range from a minimum of zero (that represents the lower bound of the index) to a maximum of $+\infty$ (that represents the upper bound of the index). The case of negative trade costs, although extraordinary and not economically meaningful, may occur anyhow.

The Heckman (1979) model performed in the entire econometric analysis takes the following form:

Selection model

$$
\begin{aligned}
\mathrm{t_{ij}} = {} & \alpha_0 + \alpha_1 \ln Distance_{ij} + \alpha_2 \ln Ruggedness_i + \alpha_3 \ln Ruggedness_j \\
& + \alpha_4 \ln Avg.Temperature_i + \alpha_5 \ln Avg.Temperature_j \\
& + \alpha_6 \ln Dist.Equator_i + \alpha_7 \ln Dist.Equator_j + \alpha_8 \ln LSCI_i \\
& + \alpha_9 \ln LSCI_j + \alpha_{10} \ln LPI_i + \alpha_{11} \ln LPI_j + \alpha_{12} \ln GCI_i \\
& + \alpha_{13} \ln GCI_j + \alpha_{14} \ln Q.Roads_i + \alpha_{15} \ln Q.Roads_j \\
& + \alpha_{16} \ln Q.Ports_i + \alpha_{17} \ln Q.Ports_j + \alpha_{18} \ln Q.Air\ Infrastr._i \\
& + \alpha_{19} \ln Q.Air\ Infrastr._j + \alpha_{20} Docs\_EXP_i + \alpha_{21} Docs\_EXP_j \\
& + \alpha_{22} \ln Cost\_EXP_i + \alpha_{23} \ln Cost\_EXP_j + \alpha_{24} Common\_Language_{ij} \\
& + \alpha_{25} Common\_Legal\_Origin_{ij} + u_{ij}
\end{aligned}
$$

Outcome model

$$
\begin{aligned}
\ln \tau_{ij} = {} & \beta_0 + \beta_1 \ln Distance_{ij} + \beta_2 \ln Ruggedness_i + \beta_3 \ln Ruggedness_j \\
& + \beta_4 \ln Avg.Temperature_i + \beta_5 \ln Avg.Temperature_j \\
& + \beta_6 \ln Dist.Equator_i + \beta_7 \ln Dist.Equator_j + \beta_8 \ln LSCI_i \\
& + \beta_9 \ln LSCI_j + \beta_{10} \ln LPI_i + \beta_{11} \ln LPI_j + \beta_{12} \ln GCI_i \\
& + \beta_{13} \ln GCI_j + \beta_{14} \ln Q.Roads_i + \beta_{15} \ln Q.Roads_j \\
& + \beta_{16} \ln Q.Ports_i + \beta_{17} \ln Q.Ports_j + \beta_{18} \ln Q.Air\ Infrastr._i \\
& + \beta_{19} \ln Q.Air\ Infrastr._j + \beta_{20} Docs_{EXP\,i} + \beta_{21} Doc\ _{EXP\,j} \\
& + \beta_{22} \ln Cost_{EXP\,i} + \beta_{23} \ln Cost_{EXP\,j} + \beta_{24} \mathrm{aht}(\rho) + \beta_{25} \ln(\sigma_u) + \varepsilon_{ij}
\end{aligned}
$$

The only regressor that is measured at the bilateral level is the weighted distance between country i and country j.[35] The other variables expressed at the bilateral level are represented by the exclusion restrictions.

In the context of trade costs, the exclusion restriction is satisfied by variables that affect the fixed cost of trade, but not the variable cost of trade. In the gravity literature, the most frequent variables exploited as latent variables have been common language (Helpman et al., 2008, Martin and Pham, 2015), common religion (Helpman et al., 2008),[36] having for long

[35] The source of the variable is the gravity data set from CEPII.
[36] Helpman, Melitz and Rubinstein (2008) exploit also the regulation cost of firm's entry.

time positive trade with the partner (Bouet et al., 2008), doing business indicators, like starting-business procedures or starting-business time (Martin and Pham, 2008).[37] Linders and de Groot (2006) and Haq et al. (2011) include the same regressors in both selection and outcome models: the assumption of normality for the error term in both regressions, and the correspondingly zero covariance, is used as identification condition.

Starting from the evidence coming from previous contributions and exploiting the information provided by the gravity data set from CEPII, two possible latent variables have been selected and included in the selection model: having the same primary official language and having the same legal origin. Beyond the link with the preceding literature, the motivation that has driven the choice of the potential excluded variables is based on the need to rely on two different backgrounds. Sharing the same language discloses the benefits that originate more from cultural concerns. Alternatively, the same legal origin reflects more bilateral trade advantages deriving from historical issues and institutions. These two variables are prone to affect the probability to engage trade with other countries but not to (less) affect variable trade costs. Other suitable potential latent variables, like having a colonial relationship or belonging to the same empire, have also been undertaken, but they resulted not relevant.

Given the above considerations, the empirical investigation proceeds in two steps. It concentrates first on the whole sample, disregarding in this stage the geography of countries. In a second phase, it looks at single geographical categories and performs separate models. In each model, at least one of the two countries of the pair belongs to the geographical category which the model refers to. Since the limitations yield by the LPI, the analysis looks at two years: 2007 (the starting year) and 2014 (the final year).[38] Moreover, since the LSCI does not exist for the group of landlocked countries, the model for the whole sample is performed both with and without the LSCI. In the model related to the single landlocked countries group, instead, the LSCI is missing. Finally, in order to provide more robust analysis, the model is estimated relying on both Maximum Likelihood and Two-Step estimators.

4.5 Estimation results

This paragraph has a twofold scope. It has the goal to present the results coming from the empirical estimations and to discover whether the analysis is able to reply to the initial

[37] The work of Martin and Pham of 2008 is the preliminary version of the subsequent work of 2015.

[38] Estimations for 2010 and 2012 have also been performed. For parsimony and since not significant differences occur across years, estimation results for the whole sample are displayed for both initial (2007) and final (2014) year, whereas estimation results by geographical category are described and reported only for 2014.

question on what determines trade costs and whether these determinants vary according to different geographical dimensions of countries.

The empirical investigation, based on the Heckman model, has been structured in two steps, investigating first the determinants of trade costs for the whole sample and, in a second stage, for the four geographical (insular) categories.

Table 4.9 shows the estimation results of the Heckman model for the whole sample for both initial (2007) and final (2014) year and for both estimators (Maximum Likelihood and Two-Step). Both results for the outcome and the selection regressions are displayed. Coefficients cannot be directly interpreted as marginal effects (or elasticities) on the dependent variable, due a change in the independent variables. This represents one of the limitations of the Heckman model that is caused by the inclusion in the outcome regression of the inverse Mills ratio. The coefficient of a given regressor in the outcome model can be interpreted as marginal effect only if that regressor is not included in the selection model. Otherwise, the parameter can be divided into two parts. One part describes the effect of a change in the variable on the probability to trade. The second part of the coefficient measures the effect of a change in the regressor on the extent of trade costs.[39] On these bases and according to the aim of this investigation, Table 4.10 displays the marginal effects for the expected value of the dependent variable conditional on being observed[40] (the intensive margin of trade) and the marginal effects for the probability of the dependent variable being observed (the extensive margin of trade). First variables in Table 4.9 and Table 4.10 refer to the geographical sphere. The distance between countries is, in each specification, highly significant and with the expected positive sign when referring to trade costs (outcome regression), and with negative

[39] Following Sigelman and Zeng (1999), the effect provided by a given variable x obtained from the Heckman estimation is equal to $\frac{\partial E(y_i|y_i^*>0,x)}{\partial x_k} = \beta_k - \gamma_k \left(\frac{\rho\sigma_\epsilon}{\sigma_u}\right) \delta_i(\alpha_u)$, where β_k measures the effect of x on $y_i(\tau_{ij}$in current analysis, the extent of trade costs) and where the latter term describes the effect of x on y_i^* (t_{ij} in current analysis, the probability to trade). $\delta_i(\alpha_u) = [\lambda_i(\alpha_u)]^2 - \alpha_u\lambda_i(\alpha_u)^4$, where λ_i is the inverse Mills ratio; $\rho = corr(u_{i,}\epsilon_i)$; σ_ϵ and σ_u represent the standard error of the residuals of the outcome and selection equation, respectively.

[40] Part of the literature on gravity model computes conditional and unconditional marginal effects, where the latter represent another dimension of the intensive margin. Conditional or unconditional marginal effects depend on the assumption of zero trade flows. If zeros in the data set are real zeros, then conditional elasticities are more appropriate. If the nature of zeros is due to misreported data, then unconditional elasticities are more suitable. Greene (2008) and Hoffmann and Kassouf (2005) derive the conditional and unconditional marginal effects. Hoffmann and Kassouf (2005) express the conditional marginal effect as $\frac{\partial E(W_i|L_i^*>0)}{\partial x_{ki}} = \beta_k - \frac{\gamma_k}{\sigma_u}\beta_\lambda\delta_i$. The expression represents the marginal effect of x_{ki} on the conditional expected value of W_i(in the work of Hoffmann and Kassouf W_i represents the earnings received by each individual). Since the conditional marginal effect measures the effect of x_{ki} on the value of W_i only for those participating in the labour market (i.e. L_i^*), the latter term does not describe the change of x_{ki} on L_i^*. The unconditional marginal effect measures, instead, the marginal effect of x_{ki} linked to a change in wage for those employed, and the effect linked to a change in the probability of being employed.

Table 4.9 Heckman model estimation results for the whole sample with and without Linear Shipping Connectivity Index (LSCI), year 2007 and year 2014

Dependent variable: Average trade costs (ln), σ=8	Maximum likelihood without LSCI				Two step without LSCI				Maximum likelihood with LSCI				Two step with LSCI			
	(1) 2007		(2) 2014		(3) 2007		(4) 2014		(5) 2007		(6) 2014		(7) 2007		(8) 2014	
	Outcome	Selection	Outcome	Selection	Outcome	Selection	Outcome	Selection	Outcome	Selection	Outcome	Selection	Outcome	Selection	Outcome	Selection
Weighted Distance i-j (ln)	0.320***	-0.399***	0.332***	-0.258***	0.306***	-0.407***	0.326***	-0.257***	0.335***	-0.206***	0.328***	-0.093**	0.322***	-0.202***	0.327***	-0.093**
Ruggedness of i (ln)	0.017**	0.081**	0.016***	-0.075***	0.021***	0.085**	0.016**	-0.072***	0.009	0.041	0.027***	-0.030	0.016	0.047	0.028***	-0.028
Ruggedness of j (ln)	0.016***	0.0861***	0.009*	-0.174***	0.021***	0.083***	0.006	-0.177***	0.008	0.0743**	0.016***	-0.228***	0.017	0.076**	0.014**	-0.226***
Average temperature of i (ln)	0.006***	0.032***	0.006***	-0.023***	0.008***	0.030***	0.005***	-0.023***	0.009***	-0.006	0.007***	-0.039***	0.009***	-0.008	0.006***	-0.039***
Average temperature of j (ln)	0.007***	0.034***	0.004**	-0.046***	0.009***	0.032***	0.003	-0.047***	0.010***	-0.009	0.004*	-0.101***	0.010***	-0.015	0.002	-0.101***
Distance from Equator of i (ln)	0.842***	2.637***	0.788***	-0.902***	0.965***	2.395***	0.756***	-0.985***	0.758***	2.468***	0.735***	-1.924***	0.944***	2.327***	0.701***	-1.960***
Distance from Equator of j (ln)	0.919***	2.931***	0.663***	-1.103***	1.091***	2.694***	0.642***	-1.173***	1.015***	2.443***	0.656***	-2.484***	1.211***	2.262***	0.615***	-2.510***
Linear Shipping Connectivity Index of i (ln)									-0.124***	0.279***	-0.098***	0.064	-0.102***	0.299***	-0.097***	0.063
Linear Shipping Connectivity Index of j (ln)									-0.130***	0.357***	-0.119***	0.138***	-0.106***	0.376***	-0.117***	0.134***
Logistic Performance Index of i (ln)	-0.733***	1.997***	-1.500***	2.207***	-0.641***	2.142***	-1.428***	2.278***	-0.685***	1.102***	-1.381***	1.779***	-0.617***	1.057**	-1.349***	1.809***
Logistic Performance Index of j (ln)	-0.708***	1.635***	-1.648***	2.139***	-0.634***	1.684***	-1.580***	2.167***	-0.679***	0.814**	-1.526***	0.388	-0.647***	0.637	-1.533***	0.378
Global Competitiveness Index of i (ln)	-1.027***	0.687	-0.568***	1.372***	-0.971***	0.736*	-0.506***	1.409***	-0.726***	-0.738	-0.485***	0.454	-0.693***	-0.614	-0.461***	0.471
Global Competitiveness Index of j (ln)	-1.131***	0.904**	-0.210**	0.836**	-1.090***	0.857*	-0.180**	0.793**	-1.266***	-0.312	-0.194*	1.474***	-1.264***	-0.439	-0.145	1.497***
Quality of roads of i (ln)	0.039	-0.687***	0.350***	-0.264**	0.004	-0.763***	0.341***	-0.258**	0.066	-0.611**	0.264***	1.221***	0.008	-0.671**	0.290***	1.221***
Quality of roads of j (ln)	0.001	-0.506***	0.243***	-0.144	-0.019	-0.605***	0.239***	-0.106	0.172***	0.217	0.242***	0.211	0.186**	0.230	0.251***	0.217
Quality of port infrastructure of i (ln)	-0.141***	-0.145	-0.070**	0.423***	-0.154***	-0.117	-0.051	0.421***	-0.111**	-0.640**	-0.052	0.284	-0.179**	-0.650**	-0.041	0.283
Quality of port infrastructure of j (ln)	0.077**	-0.230*	0.072	0.232	0.057	-0.221	0.082*	0.215	-0.055	-1.271***	0.099	-0.819**	-0.152*	-1.338***	0.079	-0.809**
Quality of air infrastructure of i (ln)	0.224***	0.939***	0.054	-0.275	0.274***	0.986***	0.039	-0.304*	0.309***	1.281***	0.320***	-1.660***	0.406***	1.272***	0.279***	-1.671***
Quality of air infrastructure of j (ln)	0.196***	0.575***	0.054	0.210	0.225***	0.701***	0.065	0.210	0.387***	0.422	0.145**	1.204***	0.418***	0.601*	0.178**	1.196***
Number of documents to export of i	-0.011***	-0.001	0.009**	-0.040***	-0.012***	-0.003	0.007*	-0.041***	-0.007	0.040	0.0148***	-0.100***	-0.003	0.035	0.013**	-0.100***
Number of documents to export of j	0.001	-0.0258*	0.027***	-0.005	-0.001	-0.025*	0.028***	-0.004	-0.009*	-0.021	0.028***	0.108***	-0.012	-0.028	0.032***	0.109***
Cost to export (US$ per container) of i (ln)	-0.014	-0.254***	0.077***	-0.105**	-0.020	-0.230***	0.076***	-0.099*	0.027	0.047	0.055***	0.348***	0.041	0.116	0.064***	0.351***
Cost to export (US$ per container) of j (ln)	0.014	-0.185**	0.079***	-0.404***	0.006	-0.208***	0.061***	-0.410***	0.0508**	-0.002	0.019	-0.803***	0.054	0.062	0.001	-0.799***
Common language i-j		0.909***		0.305***		0.849***		0.279***		1.021***		0.194**		0.862***		0.175*
Common legal origin i-j		0.172***		0.184***		0.104*		0.170***		-0.059		0.134**		-0.068		0.135**
Constant	1.274***	-1.323	-0.359	0.744	0.961***	-1.297	-0.465*	0.718	0.388	-0.210	-0.013	2.694*	-0.031	-0.745	-0.079	2.613*
Total observations	4656		6105		4656		6105		2628		3486		2628		3486	
Censored observations	894		1538		894		1538		289		621		289		621	
Log-likelihood	-2842.524		-4083.453						-1209.312		-1870.274					
Wald chi-squared statistic (probability)	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	
Aht(ρ)	0.520***		0.163***						0.427***		0.064					
Ln(σ)	-1.041***		-1.046***						-1.176***		-1.176***					
Inverse Mills ratio (λ)					0.386***		0.185***						0.499***		0.107**	

Note: Asterisks denote significance levels; * p<0.10, ** p<0.05 and *** p<0.01. P-values based on clustered standard errors

Table 4.10 Marginal effects from Heckman model estimation for the whole sample with and without Linear Shipping Connectivity Index (LSCI), year 2007 and year 2014

Dependent variable: Average trade costs (ln), σ=8	Maximum likelihood without LSCI				Two step without LSCI				Maximum likelihood with LSCI				Two step with LSCI			
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
	2007		2014		2007		2014		2007		2014		2007		2014	
	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)
Weighted Distance i-j (ln)	0.339***	-0.065***	0.337***	-0.063***	0.349***	-0.066***	0.343***	-0.062***	0.339***	-0.021***	0.328***	-0.016**	0.341***	-0.020***	0.329***	-0.016**
Ruggedness of i (ln)	0.013**	0.013**	0.018***	-0.018***	0.012	0.014**	0.021***	-0.017***	0.008	0.004	0.028***	-0.005	0.012	0.005	0.029***	-0.005
Ruggedness of j (ln)	0.012**	0.014***	0.013**	-0.042***	0.012*	0.014***	0.019***	-0.043***	0.006	0.008*	0.017***	-0.039***	0.010	0.008**	0.021***	-0.039***
Average temperature of i (ln)	0.004**	0.005***	0.006***	-0.006***	0.004**	0.005***	0.007***	-0.006***	0.009***	-0.001	0.007***	-0.007***	0.010***	-0.001	0.008***	-0.007***
Average temperature of j (ln)	0.005***	0.006***	0.005***	-0.011***	0.006***	0.005***	0.006***	-0.011***	0.010***	-0.001	0.004**	-0.017***	0.012***	-0.002	0.005**	-0.017***
Distance from Equator of i (ln)	0.720***	0.432***	0.807***	-0.219***	0.714***	0.391***	0.824***	-0.239***	0.701***	0.252***	0.746***	-0.331***	0.727***	0.235***	0.761***	-0.337***
Distance from Equator of j (ln)	0.784***	0.480***	0.686***	-0.268***	0.809***	0.439***	0.722***	-0.285***	0.958***	0.250***	0.670***	-0.427***	0.999***	0.228***	0.691***	-0.432***
Linear Shipping Connectivity Index of i (ln)									-1.131***	0.029***	-0.098***	0.011	-0.130***	0.030***	-0.099***	0.011
Linear Shipping Connectivity Index of j (ln)									-1.139***	0.037***	-0.120***	0.024***	-0.141***	0.038***	-0.121***	0.023***
Logistic Performance Index of i (ln)	-0.825***	0.327***	-1.546***	0.536***	-0.866***	0.349***	-1.584***	0.533***	-0.711***	0.113***	-1.139***	0.306***	-0.716***	0.107**	-1.404***	0.311***
Logistic Performance Index of j (ln)	-0.783***	0.268***	-1.693***	0.520***	-0.811***	0.275***	-1.729***	0.526***	-0.698***	0.083**	-1.528***	0.067	-0.707***	0.064*	-1.544***	0.065
Global Competitiveness Index of i (ln)	-1.058***	0.113	-0.597***	0.333***	-1.048***	0.120*	-0.602***	0.342***	-0.708***	-0.075	-0.487***	0.078	-0.636***	-0.062	-0.476***	0.081
Global Competitiveness Index of j (ln)	-1.173***	0.148**	-0.227***	0.203**	-0.180***	0.140*	-0.235***	0.192**	-1.259***	-0.032	-0.202**	0.254***	-1.223***	-0.044	-0.190*	0.258***
Quality of roads of i (ln)	0.071**	-0.113***	0.356***	-0.064**	0.084*	-0.124***	0.359***	-0.063**	0.081*	-0.062**	0.257***	0.210***	0.071	-0.068**	0.253***	0.210***
Quality of roads of j (ln)	0.025	-0.083***	0.246***	-0.035	0.045	-0.099***	0.246***	-0.026	0.167***	0.022	0.241***	0.036	0.165**	0.023	0.244***	0.037
Quality of port infrastructure of i (ln)	-0.134***	-0.024	-0.079**	0.103***	-0.141***	-0.019	-0.079**	0.102***	-0.096**	-0.065**	-0.053	0.049	-0.118	-0.066**	-0.050	0.049
Quality of port infrastructure of j (ln)	0.088**	-0.038*	0.067	0.056	0.080**	-0.036	0.067	0.052	-0.025	-0.130***	0.104	-0.141**	-0.027	-0.135***	0.104	-0.139**
Quality of air infrastructure of i (ln)	0.180***	0.154***	0.060	-0.067	0.171***	0.161***	0.060	-0.074*	0.278***	0.131***	0.329***	-0.286***	0.287**	0.129***	0.330***	-0.288***
Quality of air infrastructure of j (ln)	0.170***	0.094***	0.049	0.051	0.152***	0.114***	0.050	0.051	0.378***	0.043	0.139**	0.207***	0.362***	0.061*	0.142*	0.206***
Number of documents to export of i	-0.011***	-0.000	0.010***	-0.010***	-0.011**	-0.000	0.010**	-0.010***	-0.008	0.004	0.015***	-0.017***	-0.006	0.003	0.016***	-0.017***
Number of documents to export of j	0.002	-0.004*	0.027***	-0.001	0.002	-0.004*	0.028***	-0.001	-0.009*	-0.002	0.027***	0.018***	-0.009	-0.003	0.028***	0.019***
Cost to export (US$ per container) of i (ln)	-0.002	-0.042***	0.079***	-0.026**	0.004	-0.037***	0.082***	-0.024**	0.026	0.005	0.053***	0.060***	0.031	0.012	0.053***	0.060***
Cost to export (US$ per container) of j (ln)	0.022	-0.030**	0.087***	-0.098***	0.027	-0.034*	0.089***	-0.100***	0.051**	-0.000	0.023	-0.138***	0.049	0.006	0.025	-0.138***
Common language i-j		0.093***		0.065***		0.089***		0.006***		0.059***		0.030**		0.053***		0.028**
Common legal origin i-j		0.027***		0.043***		0.017*		0.004***		-0.006		0.022**		-0.007		0.023**

Note: Asterisks denote significance levels; * p<0.10, ** p<0.05 and *** p<0.01. P-values based on clustered standard errors

sign when referring to the participation to trade (selection regression). Bilateral distance reduces international trade between countries and makes trade costs higher. Ruggedness and average temperature have, as expected, a positive effect on trade costs, since they make trade more costly, and they induce countries to trade less internationally. However, in those specifications referring to 2007, the effect on the probability to trade is positive (or not statistically significant) rather than negative. This positive effect is less obvious. Having a higher fraction of the land that is rugged or having a higher average temperature represent a natural impediment and should affect negatively the participation to trade. In 2014, the expected negative sign is however confirmed. Analogously, the effect of the latitude is less clear. The distance from Equator should represent a big obstacle in taking part to international exchanges: the coefficient of the variable is highly significant and with the expected positive sign only in half of all selection regressions.[41] The high value of the coefficient, compared to the coefficients of the other geographical variables, informs of how being far from the Equator (and this is typically the case of high or middle-high income countries) plays an important role in being involved in global trade flows. However, the positive coefficient linked to the variable, when considering the extent of trade costs, tells that, the greater the distance from the Equator is, the higher trade costs are. These unclear signs connected to the coefficients of geographical regressors occur considering both country i and country j, both specifications, with and without the LSCI, and both estimators. However, the inclusion of the LSCI in the last four specifications makes some coefficients of geographical variables not significant anymore. When looking at what has been called the 'logistics, connectivity and competitiveness' dimension, Table 4.9 and Table 4.10 clearly show how in all eight specifications the three indexes report a highly significant coefficient, with the expected sign in both selection and outcome regressions and a big effect on both trade costs and trade participation. The highest effect on both trade costs and trade participation is played by the LPI and the GCI; the LSCI performs, compared to the other two indexes, a lower impact. When considering the infrastructure, opposite signs and non-significant coefficient occur more frequently. Looking at the road quality, results suggest that the quality of roads induces countries to not trade internationally or makes trade more costly. However, in specifications 6 and 8 (that refer to 2014) the effect of the probability to trade is correctly positive. The quality of ports regressor performs better than the quality of roads variable: a higher quality of ports has a negative effect on trade costs, reducing them, and encourages countries to trade more

[41] Those that refer to 2007. In selection regressions referring to 2014 the effect is negative.

with other international partners. However, also for the quality of ports inconsistent signs occur in some regressions. A different result emerges when considering the quality of air infrastructure. Having a higher quality of air infrastructure represents an important incentive to participate to global exchanges, but the effect on trade costs is positive rather negative, making them more costly.[42] The number of documents needed to export has a minor effect on both participation to international exchanges and the extent of trade costs. The expected signs are confirmed and the frequency of opposite signs is very low. Accordingly, the cost to export (US$ per container) has in both outcome and selection regressions the expected sign: it influences positively the size of trade costs, making them higher, and negatively impacts on the participation to trade. At the bottom of Table 4.9 and Table 4.10 are included the exclusion restrictions exploited. They are relevant in all specifications. But, in the fifth and in the seventh specification, the common legal origin seems to not affect trade, and only having the same language persists.

Further elements of the Heckman econometric analysis are represented by ρ, σ and λ, which provide information on the relationship between the outcome and the selection equations.[43] ρ is a likelihood ratio representing the correlation between the error terms of the outcome and the selection regressions. If the two error terms are correlated, then the sample selection, typical of this kind of analysis, generates bias. When the parameter ρ is statistically significant and the absolute value is large (the lower limit is 0, the upper limit is 1), it means that sample selection should be taken seriously, since it may create bias in the analysis. As shown in Table 4.9, the parameter ρ (aht(ρ)) is highly significant and the absolute value is large in both specifications 1 and 5 and moderate in specification 2, suggesting that sample selection can create considerable bias.[44] σ (ln(σ)) represents the standard error of the residuals of the selection equation. λ is the inverse Mills ratio obtained as ρ*σ. It represents an additional explanatory variable included in the outcome model, which controls for the influence of unobserved characteristics of the variables. The inverse Mills ratio is statistically significant in all specifications, suggesting that sample selection bias is a potential problem with the sample.

The general impression coming from Table 4.9 and Table 4.10 is that the Maximum Likelihood and the Two-step estimators produce similar results in terms of significance, signs

[42] Some opposite signs arise also for the selection regressions.

[43] The Maximum Likelihood estimator of the Heckman model does not provide ρ and σ, but provides the arc hyperbolic tangent of ρ and the natural logarithm of σ.

[44] The statistical software Stata, used for the entire empirical analysis, also provides a likelihood ratio test for the null hypothesis that ρ=0, i.e. that the two error terms are uncorrelated. Although not shown in Table 4.9, the test leads to reject the null hypothesis, suggesting that some bias may rise due sample selection.

and magnitude of effects. 2014 benefits from more observations than 2007 and, therefore, more data variability and better results. Although unclear effects produced by the geography and the infrastructure, the evidence is that the distance between countries, the overall performance, the economic competitiveness and the maritime connectivity matter for the participation in international exchanges and are important determinants of the extent of trade costs. Factors connected to the border compliance or to costs of exports reveal to be less important in explaining participation to trade and the extent of trade costs.

When looking at the different geographies of countries, the effects of the potential determinants of trade costs may change according to the geographical group they refer to. Since the inference analysis on the whole sample has not displayed significant differences between initial and final year and since the higher variability in the final year allows to achieve better estimates, estimations by geographical category are presented and discussed only for 2014.[45] Table 4.11 and Table 4.12 include, respectively, the estimation results and the corresponding marginal effects by geographical category, where at least one of the two partners belongs to that category. They allow to compare all insular categories, since the LSCI is not included within the explanatory variables.

Analogously to the estimations for the whole sample, the first set of variables listed in Table 4.11 and Table 4.12 refer to the geographical sphere. Bilateral distance is an important determinant for both trade costs and trade participation in all geographical categories, even if for the partial-insularity and for the island-states group distance it seems to not affect the probability to trade. The effect of ruggedness and average temperature is low and with the expected sign; only in few cases they report a not statistically significant coefficient or an opposite sign, and this occurs mainly for the partial-insularity category. It should be considered that the estimates for the partial-insularity group refer to a reduced sample and to the trade with just high or middle-high income countries, suggesting that a low variability occurs. Once again, the distance from the Equator does not perform properly: it shows a coefficient with the correct sign only for the selection model of the partial-insularity category.

When considering the 'logistics, connectivity and competitiveness' dimension, both LPI and GCI perform in the right way in all geographical groups. The sphere linked to infrastructures displays some weaknesses. The quality of roads reports always an opposite sign. It seems that having a better quality of roads does not induce countries to take part to trade or does not have a beneficial effect on trade costs. A correct sign is reported only for the selection equation of

[45] Although not displayed, estimations have also been performed for 2007, 2010 and 2012.

Table 4.11 Heckman model estimation results by geographical category without LSCI, year 2014

	Landlocked				Coastal				Partial-insularity				Island-states			
Dependent variable: Average trade costs (ln), σ=8	Maximum likelihood		Two step		Maximum likelihood		Two step		Maximum likelihood		Two step		Maximum likelihood		Two step	
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
	Outcome	Selection	Outcome	Selection	Outcome	Selection	Outcome	Selection	Outcome	Selection	Outcome	Selection	Outcome	Selection	Outcome	Selection
Weighted Distance i-j (ln)	0.350***	-0.558***	0.346***	-0.562***	0.311***	-0.296***	0.300***	-0.304***	0.326***	0.017	0.331***	0.008	0.312***	-0.058	0.312***	-0.059
Ruggedness of i (ln)	0.041***	-0.114***	0.040***	-0.115***	0.020***	0.059***	0.024***	0.060***	0.029***	0.126**	0.046***	0.134***	0.050***	-0.194***	0.050***	-0.192***
Ruggedness of j (ln)	0.012**	-0.039	0.012*	-0.042	0.011***	-0.055***	0.011***	-0.060***	0.020***	0.008	0.023*	0.009	0.010**	-0.194***	0.010**	-0.192***
Average temperature of i (ln)	0.005**	-0.033***	0.004**	-0.033***	-0.002	-0.011**	-0.003*	-0.012**	0.012***	0.016*	0.014***	0.014*	0.016***	-0.025**	0.016***	-0.024**
Average temperature of j (ln)	0.008***	-0.020***	0.008***	-0.022***	0.004***	-0.029***	0.004***	-0.032***	0.011***	-0.000	0.011***	-0.001	0.002	-0.056***	0.002	-0.056***
Distance from Equator of i (ln)	0.405***	-0.887***	0.390***	-0.978***	0.391***	-0.001	0.376***	-0.161	1.372***	2.559***	1.703***	2.347***	1.061***	-0.949**	1.062***	-0.986**
Distance from Equator of j (ln)	0.484***	0.382	0.492***	0.240	0.579***	-0.453**	0.567***	-0.662***	0.884***	0.874**	0.993***	0.762*	0.533***	-2.170***	0.530***	-2.210***
Logistic Performance Index of i (ln)	-0.747***	2.076***	-0.731***	2.200***	-1.173***	1.863***	-1.100***	1.982***	-0.690***	1.373**	-0.574***	1.290**	-0.695***	0.583*	-0.694***	0.615*
Logistic Performance Index of j (ln)	-0.926***	3.089***	-0.898***	3.162***	-1.143***	2.201***	-1.057***	2.271***	-1.104***	0.671*	-1.048***	0.693*	-1.256***	1.958***	-1.251***	1.981***
Global Competitiveness Index of i (ln)	-0.588***	1.061***	-0.573***	1.136***	-0.383***	1.565***	-0.306***	1.623***	-0.584***	-0.798	-0.720***	-0.710	-0.958***	0.655	-0.954***	0.689
Global Competitiveness Index of j (ln)	-0.410***	0.310	-0.408***	0.243	-0.404***	0.363	-0.393***	0.329	-0.404***	0.248	-0.408**	0.124	-0.476***	0.286	-0.476***	0.316
Quality of roads of i (ln)	0.266***	-0.508***	0.261***	-0.556***	0.101***	-0.737***	0.069***	-0.720***	0.061*	-0.753***	0.018	-0.715***	0.399***	0.511**	0.401***	0.501**
Quality of roads of j (ln)	0.186***	-0.798***	0.179***	-0.770***	0.142***	-0.676***	0.111***	-0.640***	0.097***	-0.366**	0.050	-0.308*	0.096***	0.059	0.096***	0.062
Quality of port infrastructure of i (ln)	-0.113***	0.480***	-0.108***	0.474***	-0.113***	0.362***	-0.094***	0.348***	-0.011	-0.118	-0.001	0.072	0.158***	-1.301***	0.153***	-1.296***
Quality of port infrastructure of j (ln)	-0.003	0.198	-0.003	0.207	0.050*	0.206**	0.059**	0.196**	0.001	0.184	0.023	0.182	0.009	-0.291*	0.008	-0.288*
Quality of air infrastructure of i (ln)	-0.040	0.649***	-0.033	0.671***	0.057*	0.228*	0.059*	0.198*	-0.053	1.602***	0.127	1.497***	-0.266***	0.961***	-0.263***	0.954***
Quality of air infrastructure of j (ln)	-0.025	0.806***	-0.015	0.771***	-0.048	0.906***	-0.003	0.897***	-0.015	0.974***	0.133	0.974***	-0.036	0.185	-0.036	0.176
Number of documents to export of i	0.000	-0.029**	0.000	-0.030**	0.010***	-0.015	0.010***	-0.017*	0.066***	0.006	0.063***	0.005	0.031***	-0.138***	0.031***	-0.139***
Number of documents to export of j	0.008*	-0.037**	0.008*	-0.038***	-0.003	-0.024**	-0.004	-0.023**	0.004	-0.0545**	-0.004	-0.049**	0.008*	-0.114***	0.008*	-0.114***
Cost to export (US$ per container) of i (ln)	0.093***	-0.437***	0.089***	-0.420***	0.007	-0.123***	0.004	-0.106***	0.017	0.086	0.043	0.178	0.212***	-0.554***	0.210***	-0.542***
Cost to export (US$ per container) of j (ln)	0.096***	-0.316***	0.093***	-0.320***	0.105***	-0.393***	0.085***	-0.393***	0.147***	-0.064	0.135***	-0.077	0.150***	-0.481***	0.148***	-0.473***
Common language i-j		0.617***		0.598***		0.518***		0.458***		0.525***		0.375***		0.325***		0.292***
Common legal origin i-j		0.257***		0.215***		0.117***		0.106***		0.054		-0.009		0.055		0.051
Constant	-1.129***	4.024***	-1.143***	3.914***	0.335**	0.648	0.254	0.577	-1.368***	-3.919**	-2.176***	-4.373***	-1.624***	8.301***	-1.613***	8.069***
Total observations	5833		5833		12393		12393		3912		3912		3903		3903	
Censored observations	1857		1857		2639		2639		324		324		689		689	
Log-likelihood	-3907.921				-8520.616				-1624.952				-1947.181			
Wald chi-squared statistic (probability)	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000	
Aht(ρ)	0.293***				0.272***				0.289***				0.080			
Ln(σ)	-0.986***				-1.004***				-1.205***				-1.222***			
Inverse Mills ratio (λ)			0.141***				0.249***				0.742***				0.036	

Note: Asterisks denote significance levels; * p<0.10, ** p<0.05 and *** p<0.01. P-values based on clustered standard errors

Table 4.12 Marginal effects from Heckman model estimation by geographical category without LSCI, year 2014

Dependent variable: Average trade costs (ln), σ=8	Landlocked				Coastal				Partial-insularity				Island-states			
	Maximum likelihood		Two step		Maximum likelihood		Two step		Maximum likelihood		Two step		Maximum likelihood		Two step	
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)
Weighted Distance i-j (ln)	0.374***	-0.154***	0.379***	-0.155***	0.321***	-0.061***	0.325***	-0.063***	0.326***	0.002	0.330***	0.001	0.312***	-0.010	0.313***	-0.010
Ruggedness of i (ln)	0.046***	-0.032***	0.046***	-0.032***	0.018***	0.012***	0.019***	0.012***	0.027***	0.012**	0.029	0.012***	0.051***	-0.033***	0.052***	-0.033***
Ruggedness of j (ln)	0.014**	-0.011	0.014**	-0.011	0.013***	-0.011***	0.016***	-0.012***	0.020***	0.001	0.022	0.001	0.112**	-0.033***	0.012**	-0.033***
Average temperature of i (ln)	0.006***	-0.009***	0.006***	-0.009***	-0.002	-0.002**	-0.002	-0.002	0.012***	0.001*	0.013***	0.001*	0.016***	-0.004**	0.017***	-0.004**
Average temperature of j (ln)	0.009***	-0.005***	0.009***	-0.006***	0.005***	-0.006***	0.006***	-0.007***	0.011***	-0.000	0.011***	-0.000	0.002	-0.010***	0.002	-0.009***
Distance from Equator of i (ln)	0.444***	-0.245**	0.446***	-0.269***	0.391***	-0.000	0.389***	-0.033	1.335***	0.234***	1.400***	0.216***	1.067***	-0.162**	1.072***	-0.168**
Distance from Equator of j (ln)	0.467***	0.106	0.478***	0.006	0.593***	-0.094**	0.621***	-0.137***	0.871***	0.080**	0.895***	0.070*	0.548***	-0.371***	0.552***	-0.377***
Logistic Performance Index of i (ln)	-0.838***	0.574***	-0.859***	0.606***	-1.232***	0.387***	-1.261***	0.410***	-0.710***	0.126**	-0.741***	0.119**	-0.699***	0.100*	-0.701***	0.105*
Logistic Performance Index of j (ln)	-1.061***	0.855***	-1.081***	0.870***	-1.214***	0.457***	-1.242***	0.470***	-1.114***	0.061*	-1.138***	0.064**	-1.269***	0.335***	-1.271***	0.338***
Global Competitiveness Index of i (ln)	-0.635***	0.294***	-0.639***	0.313***	-0.433***	0.325***	-0.438***	0.336***	-0.572***	-0.073	-0.628**	-0.065	-0.962***	0.112	-0.961***	0.118
Global Competitiveness Index of j (ln)	-0.424***	0.086	-0.422***	0.067	-0.415***	0.075	-0.420***	0.068	-0.408***	0.023	-0.424**	0.011	-0.478***	0.049	-0.479***	0.054
Quality of roads of i (ln)	0.288***	-0.140***	0.293***	-0.153***	0.124***	-0.153***	0.128***	-0.149***	0.072**	-0.069***	0.110	-0.066***	0.395***	0.087***	0.396***	0.086**
Quality of roads of j (ln)	0.221***	-0.221***	0.223***	-0.212***	0.163***	-0.140***	0.163***	-0.132***	0.102***	-0.033**	0.090	-0.028*	0.095***	0.010	0.095***	0.011
Quality of port infrastructure of i (ln)	-0.134***	0.133***	-0.136***	0.131***	-0.125***	0.075***	-0.122***	0.072***	-0.009	-0.011	-0.010	0.007	0.166***	-0.222***	0.166***	-0.221***
Quality of port infrastructure of j (ln)	-0.012	0.055	-0.015	0.057	0.044*	0.043**	0.043	0.041**	-0.002	0.017	-0.000	0.017	0.011	-0.050*	0.011	-0.049*
Quality of air infrastructure of i (ln)	-0.068	0.180***	-0.071	0.185***	0.050*	0.047**	0.043	0.041*	-0.076	0.147***	-0.067	0.138***	-0.272***	0.164***	-0.272***	0.163***
Quality of air infrastructure of j (ln)	-0.061	0.223***	-0.060	0.212***	-0.076**	0.188***	-0.076**	0.186***	-0.030	0.089***	0.007	0.090***	-0.037	0.032	-0.037	0.030
Number of documents to export of i	0.001	-0.008**	0.002	-0.008***	0.011**	-0.003	0.011***	-0.004*	0.066***	0.001	0.063***	0.000	0.032***	-0.024***	0.032***	-0.024***
Number of documents to export of j	0.009**	-0.010**	0.010**	-0.010***	-0.002	-0.005**	-0.002	-0.005**	0.044	-0.005**	0.002	-0.005**	0.009**	-0.020***	0.009**	-0.019***
Cost to export (US$ per container) of i (ln)	0.112***	-0.121***	0.114***	-0.116***	0.011	-0.025***	0.013	-0.022***	0.016	0.008	0.020	0.016	0.216***	-0.095***	0.216***	-0.093***
Cost to export (US$ per container) of j (ln)	0.109***	-0.088***	0.111***	-0.088***	0.117***	-0.082***	0.117***	-0.081***	0.148***	-0.006	0.145***	-0.007	0.153***	-0.082***	0.153***	-0.081***
Common language i-j		0.134***		0.130***		0.085***		0.077***		0.034***		0.027***		0.048***		0.044***
Common legal origin i-j		0.068***		0.057***		0.024***		0.021***		0.005		-0.001		0.009		0.009

Note: Asterisks denote significance levels; * p<0.10, ** p<0.05 and *** p<0.01. P-values based on clustered standard errors

the island-states group. The quality of the port infrastructure and that of the air infrastructure perform better. Having a high-quality in access to seaports seems to be particularly important for landlocked countries, where the effect in both outcome and selection regressions is higher than for the other insular categories. Surprisingly, complete opposite signs occur for the island-states group. For the partial-insularity category the variable is not relevant. Having a better quality of air infrastructure is important in determining participation to trade and trade costs for all groups. There is only one opposite sign for the coastal countries category. For the island-states group, however, the effect on the extent of trade costs is bigger than for the other insular categories, highlighting how having good air infrastructure matters. The number of documents needed to export is a less crucial determinant for both trade costs and trade participation. However, it reports the correct sign in all specifications and it is mainly highly significant. Moreover, for the island-state group the importance of the variable is greater. Analogously, the cost to export reveals to be more important for island-states. However, it represents a considerable determinant for both trade costs and trade participation for all geographical categories. Finally, the statistically significance of both ρ and λ suggests that the bias from sample selection should be taken seriously.

The image that emerges from Table 4.11 and Table 4.12 shows that the LPI and the GCI are the factors that mainly determine the participation to trade and the extent of trade costs in all insular categories. The geographical distance between countries reveals to be a crucial determinant as well. The infrastructure, instead, plays an unclear role. The quality of roads negatively affects the decision to take part to global exchanges and makes trade more costly. The quality of port infrastructure is, above all, important for landlocked countries. The quality of air infrastructure is a key issue in determining participation to trade in all geographical groups, but the favourable effect of reducing trade costs is performed only within the island-states category. Number of documents and cost to export play a more marginal role. However, for the island-states group the effect of cost per container is particularly important in determining the size of trade costs.

Table 4.13 and Table 4.14 show, respectively, the estimation results and the marginal effects of the Heckman model analysis by geographical category for 2014, including in the model the LSCI. The inclusion of the LSCI forces to not consider the landlocked countries, for which the index is not available. The first geographical variables (bilateral distance, ruggedness and average temperature) have all the expected impact. Only in few cases the linked coefficient is not statistically significant and only in three cases it displays an incorrect sign. Among the three regressors, bilateral distance is the one that performs the biggest effect. But, it should be

Table 4.13 Heckman model estimation results by geographical category with LSCI, year 2014

	Coastal				Partial-insularity				Island-states			
Dependent variable: Average trade costs (ln), σ=8	Maximum likelihood (1)		Two step (2)		Maximum likelihood (3)		Two step (4)		Maximum likelihood (5)		Two step (6)	
	Outcome	Selection	Outcome	Selection	Outcome	Selection	Outcome	Selection	Outcome	Selection	Outcome	Selection
Weighted Distance i-j (ln)	0.314***	-0.180***	0.307***	-0.175***	0.338***	0.101	0.343***	0.105*	0.328***	-0.110**	0.328***	-0.111**
Ruggedness of i (ln)	0.029***	0.065***	0.034***	0.064***	0.024***	0.075	0.031***	0.083	0.060***	-0.150***	0.059***	-0.150***
Ruggedness of j (ln)	0.018***	-0.119***	0.016***	-0.118***	0.020***	-0.028	0.020***	-0.020	0.016***	-0.162***	0.016***	-0.161***
Average temperature of i (ln)	0.001	-0.033***	-0.001	-0.035***	0.013***	0.005	0.014***	0.004	0.018***	-0.059***	0.018***	-0.059***
Average temperature of j (ln)	0.008***	-0.072***	0.005***	-0.074***	0.018***	-0.046***	0.015***	-0.046***	0.007***	-0.112***	0.007***	-0.112***
Distance from Equator of i (ln)	0.183***	-0.578**	0.152**	-0.679**	1.319***	2.526***	1.469***	2.445***	0.906***	-2.978***	0.895***	-2.986***
Distance from Equator of j (ln)	0.596***	-1.120***	0.536***	-1.271***	1.019***	-0.124	0.996***	-0.166	0.630***	-3.448***	0.614***	-3.453***
Linear Shipping Connectivity Index of i (ln)	-0.123***	0.170***	-0.115***	0.174***	-0.048***	0.285***	-0.036**	0.278***	-0.078***	-0.073	-0.078***	-0.073
Linear Shipping Connectivity Index of j (ln)	-0.138***	0.291***	-0.125***	0.284***	-0.116***	0.128*	-0.108***	0.119*	-0.101***	0.094*	-0.101***	0.094*
Logistic Performance Index of i (ln)	-0.756***	0.925***	-0.727***	0.971***	-0.558***	0.337	-0.548***	0.319	-0.593***	0.594	-0.592***	0.608
Logistic Performance Index of j (ln)	-0.734***	0.597***	-0.729***	0.593***	-0.710***	-0.575	-0.775***	-0.587	-0.989***	0.308	-0.990***	0.312
Global Competitiveness Index of i (ln)	-0.359***	1.142***	-0.281***	1.200***	-0.881***	-0.414	-0.901***	-0.348	-0.953***	0.199	-0.948***	0.203
Global Competitiveness Index of j (ln)	-0.542***	1.154***	-0.472***	1.163***	-0.683***	0.804	-0.634***	0.805	-0.620***	1.201**	-0.610***	1.206**
Quality of roads of i (ln)	0.098***	-0.283**	0.080***	-0.261**	0.132***	-0.395	0.113**	-0.397	0.279***	1.586***	0.289***	1.579***
Quality of roads of j (ln)	0.149***	-0.505***	0.123***	-0.456***	0.126***	-0.244	0.114***	-0.208	0.109***	0.582***	0.113***	0.588***
Quality of port infrastructure of i (ln)	0.069**	0.097	0.076**	0.066	0.078	-0.63	0.057	-0.559	0.469***	-2.012***	0.458***	-2.015***
Quality of port infrastructure of j (ln)	0.164***	-0.999***	0.108***	-1.024***	0.145***	-0.725**	0.100*	-0.737**	0.137**	-2.285***	0.122*	-2.282***
Quality of air infrastructure of i (ln)	0.021	0.196	0.028	0.176	0.030	0.339	0.047	0.331	-0.369***	0.998**	-0.364***	1.007**
Quality of air infrastructure of j (ln)	-0.064*	1.647***	0.044	1.665***	-0.033	2.004***	0.104	2.018***	-0.046	1.844***	-0.034	1.844***
Number of documents to export of i	0.001	-0.010	0.000	-0.010	0.057***	-0.004	0.058***	0.001	0.026***	-0.139***	0.025***	-0.138***
Number of documents to export of j	0.001	0.039**	0.003	0.039**	0.002	-0.011	0.001	-0.008	0.001	-0.055*	0.001	-0.054*
Cost to export (US$ per container) of i (ln)	0.052***	0.070	0.059***	0.087	0.030	0.282*	0.045*	0.326**	0.142***	-0.108	0.141***	-0.101
Cost to export (US$ per container) of j (ln)	0.060***	-0.456***	0.038***	-0.441***	0.089***	-0.023	0.089***	-0.009	0.106***	-0.396***	0.104***	-0.392***
Common language i-j		0.498***		0.442***		0.418**		0.349**		0.143		0.129
Common legal origin i-j		-0.036		-0.026		-0.012		-0.044		-0.003		-0.005
Constant	-0.038	1.082	-0.169	0.822	-1.118***	-3.325	-1.450***	-3.873*	-1.042***	7.888***	-1.029***	7.769***
Total observations	8732		8732		2979		2979		2932		2932	
Censored observations	1377		1377		196		196		373		373	
Log-likelihood	-5130.247				-840.0705				-1161.909			
Wald chi-squared statistic (probability)	0.000		0.000		0.000		0.000		0.000		0.000	
Aht(ρ)	0.202***				0.142				0.039			
Ln(σ)	-1.098***				-1.321***				-1.285***			
Inverse Mills ratio (λ)			0.263***				0.338***				0.035	

Note: Asterisks denote significance levels; * p<0.10, ** p<0.05 and *** p<0.01. P-values based on clustered standard errors

Table 4.14 Marginal effects from Heckman model estimation by geographical category with LSCI, year 2014

Dependent variable: Average trade costs (ln), σ=8	Coastal				Partial-insularity				Island-states			
	Maximum likelihood (1)		Two step (2)		Maximum likelihood (3)		Two step (4)		Maximum likelihood (5)		Two step (6)	
	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)	E(y | y*>0)	(Pr y*>0)
Weighted Distance i-j (ln)	0.317***	-0.028***	0.319***	-0.027***	0.337***	0.007	0.338***	0.008*	0.329***	-0.013**	0.329***	-0.013**
Ruggedness of i (ln)	0.028***	0.010***	0.029***	0.010***	0.024***	0.005	0.026***	0.006	0.060***	-0.018***	0.060***	-0.018***
Ruggedness of j (ln)	0.020***	-0.019***	0.024***	-0.018***	0.020***	-0.002	0.021***	-0.001	0.016***	-0.020***	0.017***	-0.020***
Average temperature of i (ln)	0.001	-0.005***	0.002	-0.005***	0.013***	0.000	0.014***	0.000	0.018***	-0.007***	0.019***	-0.007***
Average temperature of j (ln)	0.009***	-0.011***	0.010***	-0.012***	0.018***	-0.003***	0.017***	-0.003***	0.008***	-0.014***	0.008***	-0.014***
Distance from Equator of i (ln)	0.193***	-0.090**	0.199***	-0.106**	1.305***	0.185***	1.350***	0.178***	0.913***	-0.361***	0.917***	-0.362***
Distance from Equator of j (ln)	0.615***	-0.174***	0.624***	-0.198***	1.019***	-0.009	1.004***	-0.012	0.638***	-0.418***	0.640***	-0.419***
Linear Shipping Connectivity Index of i (ln)	-0.126***	0.026***	-0.127***	0.027***	-0.050***	0.021***	-0.049***	0.020***	-0.077***	-0.009	-0.078***	-0.009
Linear Shipping Connectivity Index of j (ln)	-0.143***	0.045***	-0.145***	0.044***	-0.116***	0.009*	-0.114***	0.009*	-0.101***	0.011*	-0.102***	0.011*
Logistic Performance Index of i (ln)	-0.772***	0.143***	-0.795***	0.151***	-0.560***	0.025	-0.564***	0.023	-0.594***	0.072	-0.597***	0.074
Logistic Performance Index of j (ln)	-0.745***	0.093***	-0.770***	0.092***	-0.707***	-0.042	-0.746***	-0.043	-0.989***	0.037	-0.992***	0.038
Global Competitiveness Index of i (ln)	-0.379***	0.178***	-0.363***	0.187***	-0.879***	-0.030	-0.884***	-0.025	-0.953***	0.024	-0.950***	0.025
Global Competitiveness Index of j (ln)	-0.562***	0.179***	-0.552***	0.181***	-0.687***	0.058	-0.673***	0.059	-0.623***	0.146**	-0.619***	0.146**
Quality of roads of i (ln)	0.103***	-0.044**	0.098***	-0.041**	0.134***	-0.029	0.132**	-0.029	0.276***	0.192***	0.277***	0.191***
Quality of roads of j (ln)	0.158***	-0.079***	0.155***	-0.071***	0.128***	-0.018	0.124***	-0.015	0.108***	0.071***	0.108***	0.071***
Quality of port infrastructure of i (ln)	0.068**	0.015	0.072**	0.010	0.081	-0.046	0.085	-0.041	0.474***	-0.244***	0.473***	-0.244***
Quality of port infrastructure of j (ln)	0.181***	-0.155***	0.179***	-0.159***	0.149***	-0.053**	0.135**	-0.054**	0.142***	-0.277***	0.139**	-0.277***
Quality of air infrastructure of i (ln)	0.018	0.030	0.016	0.027	0.029	0.025	0.031	0.024	-0.372***	0.121**	-0.371***	0.122**
Quality of air infrastructure of j (ln)	-0.092**	0.256***	-0.071	0.259***	-0.044	0.147***	0.006	0.147***	-0.051	0.224***	-0.048	0.224***
Number of documents to export of i	0.001	-0.002	0.001	-0.002	0.057***	-0.000	0.058***	0.000	0.026***	-0.017***	0.026***	-0.017***
Number of documents to export of j	0.000	0.006**	0.000	0.006**	0.002	-0.001	0.001	-0.001	0.001	-0.007*	0.001	-0.007*
Cost to export (US$ per container) of i (ln)	0.051***	0.011	0.053***	0.014	0.029	0.021*	0.029	0.024**	0.142***	-0.013	0.142***	-0.012
Cost to export (US$ per container) of j (ln)	0.068***	-0.071***	0.068***	-0.069***	0.089***	-0.002	0.089***	-0.001	0.107***	-0.048**	0.107***	-0.048***
Common language i-j		0.060***		0.055***		0.023***		0.020***		0.016		0.015
Common legal origin i-j		-0.006		-0.004		-0.001		-0.003		-0.000		-0.001

Note: Asterisks denote significance levels; * p<0.10, ** p<0.05 and *** p<0.01. P-values based on clustered standard errors

noticed that, while bilateral distance has a large impact on the extent of trade costs, this is less the case when considering the participation to trade, for which distance is a minor determinant. The fourth geographical determinant (distance from Equator) face the same problems of previous estimations. A higher distance from the Equator seems to make trade more costly and to reduce the participation of countries to international exchanges. Only for the partial-insularity group the correct positive sign in the selection equation is preserved. The new introduced LSCI performs in the properly manner. It reveals to be a good predictor for both trade costs size and involvement in international trade. Particularly noteworthy is the effect for coastal countries, for which the impact of the extent of trade costs is bigger compared to the other geographical categories. Analogously, the LPI and the GCI confirm to be important determinants in the trade costs field. The infrastructure, instead, also when introducing the LSCI, has an unclear effect on both trade costs and trade participation. The quality of roads and the quality of ports always report an opposite expected sign in the linked parameters: a correct sign is displayed for the land infrastructure only in the selection regressions of the island-states group. Nevertheless, the quality of air infrastructure performs quite well. Coefficients, when statistically significant, have the correct sign and confirm how a good quality of the air transport network is important in making trade less costly and in increasing the involvement of countries in global trade. This is particularly true for the island-states category, for which the effect performed by the air infrastructure is higher, compared to the other geographical categories. With the inclusion of the maritime connectivity, the effects of the number of documents needed to export and of the cost to export loose significance and size.[46] However, although some opposite signs for the other group, the cost per container reveals to be a main predictor for the extent of trade costs within island-states. Finally, the value and significance of ρ and of the inverse Mills ratio confirm that, for the majority of the specifications, some bias due sample selection occurs.

The estimation, enhanced with the inclusion of the LSCI, has confirmed how, for all geographical categories, logistic abilities and competitiveness qualities represent the main determinants for both trade costs and trade participation. The LSCI has a more marginal role, but it seems to matter more for coastal countries than for island-states or for the partial-insularity group. Bilateral distance is particularly important in determining the extent of trade costs, but its role in influencing the involvement in global exchanges is secondary. Air infrastructure reveals to be predominant in affecting trade costs within island-states. Border

[46] Also the common legal origin (the second exclusion restriction exploited in the analysis) loses significance with the introduction of the LSCI.

compliances and costs to export play a minor effect, but US$ per container affect mainly the island-state group.

4.6 Robustness checks

The main empirical analysis so far performed has highlighted important new evidence on the main determinants of trade costs and on potential differences among geographical groups. However, it has revealed weaknesses in terms of reliability of some regressors and variability within the partial-insularity group. In order to overcome these limitations, additional inferential investigation has concentrated on a different dependent variable, on different econometric models and on different elasticities of substitution.

The dependent variable exploited in the robustness checks is the 'adjusted' version of the geometric average of international trade costs between countries i and their partner j_s, grouped according to their WB income level. In other terms, trade costs are an average by nation i and by WB income group of j_s partners. The idea is to exploit, when feasible, the variability across years,[47] across countries and across the WB income categories of the partners.

Four different econometric models have been adopted, in order to find the best one able to better describe the data: fixed effects, random effects, pooled OLS, cross-section. The use of diverse specifications starts from Novy (2013), who wonders, implementing himself a fixed effects model, whether fixed effects are effectively capturing all those unobservable elements that affect trade costs, such as regulations, formal procedures and technical barriers.[48]

Finally, trade costs have been computed using a lower (7) and a higher (11) elasticity of substitution.

Table 4.15 in Appendix 4.1 displays the estimation results by geographical category for all the four models above named and excludes the LSCI, in order to allow the comparability between insular groups. Differently from previous estimations, regressions in Table 4.15 comprise only variables for country i (since countries j_s are grouped by WB income category) and include two additional variables: the number of documents needed to import and the cost to import (US$ per container). The first impression, observing Table 4.15, is that the fixed effects model is not measuring the phenomenon in the right way. Almost all variables are not relevant, just few regressors report a significant coefficient and some parameters display a wrong sign. The very high R-squared values warn about overall accuracy of the model,

[47] A four year (2007, 2010, 2012, 2014) panel data set is exploited according to the availability of the Logistic Performance Index (LPI)

[48] "*But it is unclear whether the fixed effects capture trade cost elements that are harder to observe such as red tape and technical barriers to trade [...]*" (Novy, 2013, pp. 110-111).

suggesting that potential collinearity or sample size problems are invalidating the analysis. When considering the columns 2, 6, 10 and 14, the general feeling, coming from the random effects model, is that it is capturing the issue on trade costs determinants in a better way, although it is well-known that estimates may be biased, since there is no control for omitted variables. It must also be kept in mind that all these variables are reflecting, somehow, different shadows of the same picture (performance and ability of the country), hence a potential problem of multicollinearity may arise when introducing all these indexes together.[49] While the random effects model produces more promising results, the tests and post-estimations checks are not so comforting. The Hausman test compares fixed and random effect models under the null hypothesis that individual effects are not correlated with any independent variable in the model (Hausman, 1978). If the null hypothesis of no correlation is not rejected, both within and GLS estimators are consistent, but the within estimator is inefficient; on the other hand, if the null hypothesis is rejected, the within estimator is consistent, but the GLS is inconsistent and biased (Greene, 2008). In Table 4.15, both Hausman and a overidentifying restrictions test have been performed. The latter is a test of fixed vs. random effects as well, but is based on a Sargan-Hansen statistic.[50] The P-values of the Sargan-Hansen statistics confirm that fixed effects should be included in the model and that the GLS estimator is not consistent. Although not displayed in Table 4.15, a further test, based on the Mundlak model,[51] has been run. The Wald test on the extra-coefficients (the coefficients of the individual means of the time-varying regressors) enables to prove whether the random effects model suits the data or not. Under the null hypothesis, the coefficients of the added group-means are jointly equal to zero and the basic random effects model (i.e. without group-means) performs better than the Mundlak model (i.e. the random effects model plus the group-means variables). In all three specifications, the zero P-value leads to reject the null hypothesis, confirming that group-means are relevant in the specification and that the basic random effects model is not sufficient. All these checks seem to confirm that the model should include fixed effects rather than random effects, although the fixed effects model appears to be weaker, in terms of accuracy, than the random effects model.

[49] Table 4.17 tries to avoid a similar problem including a reduced number of regressors and excluding all those variables that potentially may create a collinearity problem.

[50] Since with heteroskedastic- and cluster-robust standard errors the Hausman test cannot be performed and since some heteroskedasticity has been detected in the data, to overcome these limitations and in order to have a robust test, the test of overidentifying restrictions has been implemented. Results of the Hausman test in Table 4.15 refer to neither heteroskedastic- nor to cluster-robust standard errors and have been included only for completeness.

[51] The Mundlak model estimates a random effects model, adding group-means of the independent variable included in the regression, which change within groups (StataCorp, 2013).

Table 4.15 also reports results from a pooled OLS estimation, assuming therefore that individual effects are constant across individuals. The suitability of the pooled OLS model is checked indirectly through the F-test on the fixed effects and with the Breusch and Pagan Lagrange multiplier (LM) test on the random effects. The F-test suggests that fixed effects seem reasonable and that there is a significant increase in goodness-of-fit in the fixed effects model compared to the pooled one. Similarly, the test on random effects hints that the random effects model is better than the pooled one. To conclude, cross-section specifications are also included in Table 4.15.

Beyond this mere econometric discussion about the suitability of the models, it is interesting to consider what results originate from the regression analysis. The image that emerges from Table 4.15 is that trade costs of landlocked countries are mainly affected by the LPI, which represents the first determinant. Geography is also important, but what immediately comes out is the positive coefficient linked to distance from the Equator variable, suggesting that, the further the country is, the higher the trade costs are. It should be reminded that the majority of countries included in the landlocked group are African countries, hence the variable, that is supposed with a negative coefficient, is reflecting the distance from the Equator mainly for the African continent. More than the wrong sing of the coefficient of the distance from Equator, what draws attention is the role played by the quality of port infrastructure. For landlocked countries, the quality of ports assesses the access to seaports. The highly significant coefficients and their high values seem to confirm how having an access to the sea is determinant for trade, also for those countries that do not have an access to the sea. Another important factor is represented by the cost to import. The group of coastal countries is the more heterogeneous in terms of wealth of countries and the more numerous. Specifications from 5 to 8 are fairly concordant to find in the LPI the most important factor affecting trade costs. Geography has a marginal role, whereas the quality of air infrastructure seems to be the second factor that determines coastal countries' trade costs. The quality of air infrastructure plays a central role for the partial-insularity group too, which in some specifications overtakes the impact of the LPI. Also for this category, the distance from the Equator has the wrong sign and it is not the only variable whose coefficient has the opposite sign. The quality of ports, the GCI and other variables mark an opposite coefficient. This problem occurs less with the coastal countries group, and the reason might be attributed to the low variability within the partial insularity category[52] and to a small sample size. The drawback of wrong signs heavily

[52] 90 per cent are rich countries.

occurs also within the island-states group. Disregarding this problem, what emerges is that the coefficient of the LPI is for island-countries not significant at all. Factors determining trade costs can be retrieved in the GCI and in the quality of the air infrastructure. The estimates coming from Table 4.15 should be taken very cautiously, since the impression is that some misspecification problems are influencing the analysis.

When introducing the LSCI, as in Table 4.16, estimate outputs and the general overview of the inference analysis are more promising. Wrong signs reduce, significance improves and the models seem to capture the phenomenon in a more correct way. However, it is not possible to provide an analysis for landlocked countries, since the LSCI is not available for them. Estimates coming from Table 4.16 depict the LPI as the first determinant of coastal countries' trade costs, followed by the quality of air infrastructure and by the LSCI. For the partial-insularity group, it is the quality of air infrastructure that plays the main role, followed by the quality of roads and by the average temperature. The coefficients linked to the quality of ports have the highest values, but the sign is positive rather than negative, suggesting that the higher the quality of port is, the higher the trade costs are. The positive coefficient of the quality of ports also occur for the island-states group. Nevertheless, the most important factor that matters for trade costs is the GCI, followed by the air infrastructure and by the LSCI.

In Table 4.17, a more parsimonious model has been exploited. Besides the geographical variables and the connectivity, logistic and competitiveness measures, in Table 4.17, the quality of infrastructure is measured only by the quality of overall infrastructure and border compliances are, instead, captured only by the Trading Across Border. The problem of wrong sign has extremely improved, although not completely solved. The island-states group reports highly significant and positive coefficients for the overall infrastructure in three out of four models; in the fixed effects model, the coefficient is not statistically significant. Results confirm that, for landlocked and coastal countries, the logistic abilities represent the most important determinant. For partial-insularity countries, it is the LSCI to become the key factor. And within the island-states group the GCI persists.

In Table 4.18, only cross-section analysis for single years (2007 and 2014) have been performed. Trade costs have been computed using a lower elasticity of substitution (σ=7) and a higher one (σ=11). Moreover, the quality of infrastructure is captured by quality of both overall infrastructure and air infrastructure. Some variables, that are important in 2007, become irrelevant in 2014. For the partial-insularity group, almost none coefficient is significant. Nevertheless, the logistic abilities for both landlocked and coastal countries and the competitiveness for island-states are confirmed as the main determinant of trade costs in all four specifications. Table 4.18 says something more. For the landlocked group, the quality of infrastructure always matters, instead the costs to export and to import seem to be important only in 2007. For both coastal and island-states groups, the shipping connectivity is relevant both in 2007 and 2014. However, wrong signs occur also in Table 4.18.

Although the robustness checks give the impression of not producing so robust and reliable results and that econometric estimation problems are plaguing the investigation, some firm points of reference emerge. Logistics abilities deeply affect trade costs in landlocked and coastal countries, but not in islands-states, where, instead, it is more the competitiveness and the performance of the country that play a big role. The shipping connectivity is important in all countries and, even though the LSCI in landlocked countries is not provided, the information comes out in any case from the relevance of the quality of ports, that, in the case of landlocked countries, corresponds with the access to seaports. Differently from expectations, the maritime connectivity is not so important for island-states, where it is more the quality of the infrastructure that matters. The quality of infrastructure is also crucial, particularly, for partial-insularity countries.

4.7 Bayesian Model Averaging: a window on future research

The main empirical analysis so far performed, the robustness checks based on different analytical models and the still little evidence coming from previous literature highlight two main facts. First, investigations on the determinants of trade costs are far from having reached a full knowledge of the phenomenon. Second, the classic inferential analysis presented in Paragraph 4.5 and in Paragraph 4.6 contribute to creating an objective understanding of what makes trade costly.

However, in terms of model specification, the analysis performed has emphasised important evidence. Whereas the Heckman model in conjunction with the 'unadjusted' version of the indirect measure of trade costs seems to capture the trade costs issue quite well, the robustness analysis has, instead, highlighted some weaknesses in these terms. The fixed effects model, although hypothetically the most suitable, appears to not be able to measure in a precise manner the determinants of trade costs. The random effects model produces more reliable results; however, post-estimation checks hint that individual effects should be included in the model and should be fixed, not random. Analogously, pooled OLS models perform better than fixed effects ones, but the test on random effects suggests that the random effects model is more suitable than the pooled one. Cross-section analysis, although more precise, highlights severe disadvantage in terms of variability, since within some geographical groups heterogeneity is completely absent.

The existence of different model classes (linear regression model, generalised linear models, nonlinear models) and the selection of subsets of explanatory variables in defining a regression model should seriously be taken into account in empirical analyses. Indeed, the existence of more than one model able to provide a good fit to the data, but yielding different effects, produces ambiguity on what is the best model able to capture the phenomenon. As pointed out by Hoeting et al. (1999), to base inference on a single model, when there is no certainty about which is the most suitable one, may mitigate the effects and the predictions coming from the empirical analysis. Leamer (1978), Draper et al. (1987) and Hodges (1987) highlight how, in case of model uncertainty, part of the evidence is exploited in the model specification.

The problem of model uncertainty affects different fields of (economic) research and, in particular, all those situations where a large set of needed explicative variables can dilute the data information. The study of the determinants of trade costs seems to perfectly fit this problem. The uncertainty lies particularly in the choice of explicative variables and model specification. In light of this, the current paragraph aims to discuss the Bayesian Model Averaging (BMA) for the determinants of trade costs theme as a new path to follow in further research, for which the extensive analysis provided in this chapter represents a good basis in terms of evidence provided and understanding of the subject.

The BMA, as underlined by the seminal literature (Roberts,1965; Leamer, 1978; Dijkstra, 1988; Draper, 1995; Chatfield, 1995; and Kass and Raftery, 1995), has the big advantage to yield improvements in predictive performance over a single selected model and to not ignore the drawbacks coming from model uncertainty. However, as remarked by Hoeting et al. (1999), there are also some difficulties linked mainly to the computation of integrals,[53] to the specification of the prior distribution or to the summation of the terms in the posterior distribution.

Whereas in the study of the potential determinants of trade costs field the use of BMA is a novelty, in the growth literature, as highlighted by Crespo Cuaresma et al. (2009), there has been several attempts to address the problem of model uncertainty relying on model averaging. One of the most significant contributions is the work of Fernández et al. (2001), which, by presenting a prior structure based on little information, uses BMA to examine the

[53] Markov chain Monte Carlo methods try to overcome these types of problems.

robustness of the determinants of growth selected in the seminal paper by Sala-i-Martin (1997). Accordingly, the work of Crespo Cuaresma and co-authors exploits improper priors for parameters based on Fernández et al. (2001), to examine a large set of potential determinants of EU regions' growth. The path that will be covered by future research on the determinants of trade costs among different geographies will be in line with the work of Crespo Cuaresma, Doppelhofer and Feldkircher. It will aim at investigating a large set of potential determinants of trade costs, exploiting BMA to explore the robustness of these determinants; moreover, it will discuss the Bayesian solution of averaging over all possible models (i.e. all possible subsets of regressors) within the selected model versus averaging over models with different error structures (i.e. generalised linear models).

To implement model averaging, three main elements are needed: the marginal distribution of the data, the prior model probabilities and the posterior distribution under the considered model.

Following Crespo Cuaresma et al. (2009) and Fernández et al. (2001) and in order to provide a better predictive information, future analysis will consider n independent replications from a selected regression model[54] with an intercept α, ι_N (a n-dimensional column vector of ones), $\mathbf{Z}$ (a $n * k$ matrix that includes k explanatory variables) and a k-dimensional vector containing the parameters. This results in 2^k possible sampling models, according to whether these regressors are included in or excluded from the model. Indeed, a given model M_j, j=1,...,2^k is the result of the selection and inclusion of $0 \leq k_j \leq k$ regressors. On these bases, the model that will be exploited to analyse the robustness of the determinants of trade costs will take the following form:

$$y = \alpha\iota_N + \boldsymbol{Z}_j\beta_j + \sigma\varepsilon$$

y is a n-dimensional column vector that represents bilateral trade costs for n country pairs, $\mathbf{Z_j}$=(z_1, ..., z_j) is a $n * k_j$ matrix that includes k_j regressors (the potential determinants of trade costs) and represents a sub-matrix of Z of all relevant explanatory variables, β_j is a j-dimensional vector containing the parameters linked to the covariates included in **Zj**, σ is a

[54] "*Model averaging also avoids the problem of having to defend the choice of model, thus simplifying presentation. Indeed, model averaging results are robust to model choice (but not necessarily robust to model class, [...]*" (Hoeting et al., 1999, p.398).

scale parameter and ε[55] is the error term that follows a *n*-dimensional Normal distribution with zero mean and diagonal covariance-variance matrix ($\Sigma = \sigma \mathbf{I}_N$).

The second step is represented by the prior probabilities of both parameters space and model space. When the prior information about the plausibility of a given model is limited, the commonplace strategy is to assume that all models are equally likely plausible (Hoeting et al., 1999). However, Spiegelhalter et al. (1993) point out how predictive performance of Bayesian approaches improves when including informative priors distributions in the model. Nevertheless, Raftery et al. (1997) and Fernández et al. (1997, 2001) provide a solution in terms of prior structure for all those cases in which little prior information is available. On these bases, according to Fernández et al (2001), non-informative improper prior will be used to specify the distribution of the parameters α and σ and g-prior distribution (Zellner, 1986) for β_j under the model M_j:

$$p(\alpha, \beta_j, \sigma | M_j)$$

$$p(\alpha) \propto 1$$

$$p(\sigma) \propto \sigma^{-1}$$

$$p(\beta_j | \alpha, \sigma, M_j) f_N^{kj} (\beta_j \,|\sigma^2 (g \boldsymbol{Z}_j' \boldsymbol{Z}_j)^{-1})$$

The space of all 2^k models is defined by *M*:

$$M = \{M_j : j = 1, \ldots, 2^k\}$$

The prior probability of each model is given by:

$$P(M_j) = p_j, j = 1, \ldots, 2^k, \qquad with\, p_j > 0\, and \sum_{j=1}^{2^k} p_j = 1$$

Last step involves the computation of posterior probabilities that, in the context of model uncertainty, depend on how the prior has been specified.

The brief overview proposed here will constitute the opening of a new piece of research. Starting from this structured and planned framework, future research will rely on the BMA approach and on Markov chain Monte Carlo methods (for computations and approximations),

[55] The error term may include country fixed effects. "*The generalization of the BMA strategy here to other error structures with fixed effects is straightforward after application of the Frisch-Waugh-Lovell theorem. In a panel setting, the estimation of fixed effect models can be carried out by estimating the model proposed above using within-transformed data*" (Crespo Cuaresma et al. 2009, p. 4).

in order to provide a further insightful understanding of the determinants of trade costs by geographies of countries.

4.8 Concluding remarks

This chapter has started with a pioneering research question and promising intentions: '*What determines trade costs, and do these determinants change when the geography of countries is controlled for?*'. The incentive to explore a similar research topic comes from a triple motivation. First, trade costs matter; they are, although the steep decrease in recent years, high, and they impede that trade between nations takes place in an inexpensive way. Second, geography matters: countries' physical features determine their fortune or their misfortune and, since geography cannot be changed or improved, ad hoc policies are needed to turn adverse geography conditions into less economic weaknesses. Third, policy strategies and recommendations matter; they are fundamental to reduce trade costs and geographical disadvantages, but they might be more powerful if trade costs are investigated in conjunction with geography.

On these bases, the indirect measure of trade costs (according to the 'top-down' approach proposed by Novy, 2013) has been constructed for 188 countries in the world and for a long time-span of 20 years (from 1995 to 2014). Using the insularity data set constructed by Pinna and Licio in 2013, the 188 countries have been distinguished in four different geographical groups: 32 landlocked countries, 87 coastal countries, 17 partial-insularity countries, 52 island-states. Moreover, to provide a better overview of the trade costs theme, the analysis has been performed in two stages, relying on the advantages coming from the 'unadjusted' version of trade costs and from the Heckman model approach. In a first stage, the investigation has focused on the zero trade flows in order to consider the case of missing costs. In a second stage, the extent of trade costs has been assessed looking at what makes trade costly.

Although this encouraging and prominent prospect on the dependent variable side, some difficulties arose when considering the determinants of trade costs. In order to avoid collinearity or omitted bias problems, the selection of potential factors, affecting mainly trade costs, has identified four main areas. The first problem met has been the merge of more than seven different data sets with different countries, different years and different data organisation. Moreover, the reduced time-span available has forced to restrict the inference analysis from a long 20 years period to a period composed by only four years. However, the main difficulty has been represented by the absence of many countries in the data sets. This problem has plagued the inference analysis in terms of heterogeneity and in terms of few observations within some geographical categories. Nevertheless, both empirical and descriptive analyses set out new evidence on trade costs.

From the descriptive analysis, it is clear how having unfavourable geographic conditions determines inefficiencies in economic terms. The geographical disadvantage of landlocked countries and island-states emerges when considering both their participation to the global trade market and the extent of their trade costs. Countries that are completely surrounded by other countries or totally surrounded by the ocean face more problems in taking part to international exchanges. These difficulties, in some cases, are so high that they impede trade completely. Even when they are able to participate to global trade flows, trade costs are higher compared to those of other geographies. Indeed, island-states and landlocked countries have the highest trade costs, partial-insularity countries the lowest. These results should be considered in light of the fact that landlocked countries are mainly low or middle-low income countries, while the partial-insularity group is composed, instead, by wealthier states. Being completely surrounded by the sea, rather than having an access to the sea, leads to more costly trade: island-states' trade costs are considerably higher than those of coastal countries, even though the two groups are similar in terms of income level composition. Indeed, the group of island-states is composed by a higher fraction of rich countries, confirming even more how to be an island represent a geographical disadvantage. What is more remarkable is that, whereas landlocked countries are able to overcome their geographical disadvantage when they are rich, this is not the case for the island-states, for which being an islands represents a geographical constraint also in wealthy conditions.

From the empirical analysis clearly emerges how distance between countries and factors connected to logistic and competitiveness abilities are the main determinants of trade costs. The maritime connectivity and the cost to export (represented by the cost per container), although important and contributing to making trade more costly, are not the major determinants. The logistic abilities really matter for landlocked and coastal countries' trade costs, but are, instead, lightly less important for partial-insularity countries and island-states. In island-states, it is more the overall performance and its economic competitiveness that matter. The infrastructure has a curious effect. The quality of infrastructure is also a central determinant of trade costs and, for some groups, the air infrastructure matters more than the shipping connectivity. Indeed, the maritime connectivity is, surprisingly, not the first determinant of trade costs for island-states, whereas the access to neighbouring countries' ports is really important for landlocked countries. In countries completely surrounded by the sea, it is more the quality of the air infrastructure that matters. Border compliances and red tapes have a more reduced impact on trade costs.

Although some deficiencies and weaknesses in the empirical analysis, some preliminary and important building blocks have emerged. It has been possible to disentangle what is, at least, the primary trade costs determinant for each geographical category. It has been possible to assess that there are differences by geographies of countries. And it has been possible to examine what determinants matter less or do not matter at all. The current analysis provides insights to think about how the study on trade costs determinants might be improved, not only from the econometric point of view, but also in terms of data organisation. Moreover, the present investigation represents a good basis for further work. Future research will focus on Bayesian Model Averaging, to deal with the potential problem of model uncertainty and to better address policy recommendations and suggestions.

APPENDIX 4.1 - ROBUSTNESS CHECKS: ESTIMATION TABLES

Table 4.15 Estimation results by geographical category without Linear Shipping Connectivity Index (LSCI)

	Landlocked				Coastal				Partial-insularity				Island-states			
Dependent variable: Average trade costs (ln), σ=8	FE	RE	Pooled	Cross-section	FE	RE	Pooled	Cross-section	FE	RE	Pooled	Cross-section	FE	RE	Pooled	Cross-section
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Ruggedness (ln)		0.116***	0.091**	0.032		0.019	-0.005	0.001		0.047	0.109**	0.195***		0.0795**	0.0854**	0.158***
Average temperature (ln)		0.288***	0.344***	0.311***		0.126**	0.030	-0.004		0.097	0.124	0.17		-0.105*	-0.067	0.111
Distance from Equator (ln)		0.123***	0.143***	0.127**		-0.120***	-0.087***	-0.062*		0.144**	0.210***	0.310***		-0.0968	-0.130*	-0.245***
Logistic Performance Index (ln)	-0.396***	-0.678***	-1.178***	-1.739***	-0.177	-0.904***	-2.259***	-2.624***	-0.306	-1.014***	-1.610***	-5.108***	0.614***	0.173	-0.33	0.761
Global Competitiveness Index (ln)	0.575	-0.248	0.333	0.383	1.243**	0.050	-0.226	-0.060	1.123	0.192	1.330*	5.282***	-1.619**	-3.209***	-2.813***	-4.893***
Quality of roads (ln)	-0.092	-0.208*	-0.097	0.226	-0.080	0.076	0.223*	0.230	-0.847**	-0.737***	-1.002***	-1.571***	0.414	0.290**	0.348**	0.368*
Quality of port infrastructure (ln)	0.126	-0.384***	-0.710***	-0.926***	-0.294**	-0.155	-0.058	0.044	1.435***	1.337***	1.278**	1.25	0.126	0.967***	1.224***	3.266***
Quality of air infrastructure (ln)	0.260*	0.146	-0.062	-0.242	-0.074	-0.291**	-0.232	-0.351*	-1.134**	-1.421***	-1.712***	-1.186	-0.192	-0.607**	-0.919**	-2.651***
Number of documents to export	0.051	0.057**	0.036	0.005	0.005	0.031*	0.012	0.024	0.0842	0.180***	0.250***	0.386***	-0.0295	0.0379*	0.020	0.216***
Cost to export (US$ per container) (ln)	-0.244	-0.047	-0.138	-0.112	0.060	0.005	-0.075	-0.070	0.999	0.007	-0.529	-1.831**	0.271	0.450*	1.273***	6.090***
Number of documents to import	-0.034	-0.010	0.004	0.015	0.003	-0.008	-0.026	-0.031	-0.0996	-0.005	-0.0343	-0.147**	0.0156	0.0278	0.038	-0.0206
Cost to import (US$ per container) (ln)	0.300*	0.359***	0.369**	0.232	0.111	-0.019	-0.102	-0.085	-0.554	-0.080	0.291	1.726**	-0.633	-0.269	-0.861***	-4.772***
WB income group of js partners		0.205***	0.211***	0.200***		0.229***	0.233***	0.232***		0.241***	0.240***	0.241***		0.279***	0.279***	0.264***
Constant	0.813	-0.217	0.489	2.223**	-0.466	2.186***	5.271***	5.301***	-1.794	2.759*	3.816***	1.11	5.399*	3.181**	1.663	-3.786*
Individual fixed effects	YES	YES	NO	NO	YES	YES	NO	NO	YES	YES	NO	NO	YES	YES	NO	NO
Time fixed effects	YES	YES	YES	NO	YES	YES	YES	NO	YES	YES	YES	NO	YES	YES	YES	NO
Observations	372	356	356	104	852	852	852	256	218	202	202	56	224	224	224	64
Groups	108	104			256	256			56	52			64	64		
F-statistic overall significance	3.53	532.76	48.39	59.53	3.13	463.14	44.68	46.75	1.03	470.27	28.11	34.23	1.58	873.33	57.70	92.55
F-statistic overall significance (P-value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.436	0.000	0.000	0.000	0.122	0.000	0.000	0.000
F-test on fixed effects (P-value)	0.000				0.000				0.000				0.000			
Breusch-Pagan LM test on random effects (P-value)		0.000				0.000				0.000				0.000		
Hausman test (FE vs RE)		0.000				0.000				0.000				0.000		
Sargan-Hansen Statistic (FE vs RE)		0.000				0.000				0.000				0.000		
R-squared	0.934	0.792	0.809	0.875	0.900	0.549	0.607	0.692	0.878	0.721	0.736	0.866	0.928	0.813	0.824	0.922

Note: Asterisks denote significance levels; * $p<0.10$, ** $p<0.05$ and *** $p<0.01$. P-values based on clustered standard errors

Table 4.16 Estimation results by geographical category with Linear Shipping Connectivity Index (LSCI)

Dependent variable: Average trade costs (ln), σ=8	Coastal				Partial-insularity				Island-states			
	FE	RE	Pooled	Cross-section	FE	RE	Pooled	Cross-section	FE	RE	Pooled	Cross-section
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Ruggedness (ln)		0.028	0.017	0.044**		0.139***	0.159***	0.165**		0.096**	0.098***	0.117**
Average temperature (ln)		0.193***	0.162***	0.143***		0.398***	0.395***	0.567*		0.031	0.081	0.101
Distance from Equator (ln)		-0.101***	-0.096***	-0.085***		0.153**	0.177***	0.240***		-0.159***	-0.177***	-0.217***
Logistic Performance Index (ln)	-0.155	-0.684***	-1.379***	-1.352***	-0.223	-0.325	-0.742	-2.658	0.534***	0.163	-0.331	0.066
Linear Shipping Connectivity Index (ln)	-0.016	-0.178***	-0.216***	-0.237***	-0.051	-0.305***	-0.343***	-0.520	-0.135	-0.243***	-0.275***	-0.147
Global Competitiveness Index (ln)	1.316**	0.202	-0.100	-0.080	1.071	-0.117	0.291	1.193	-1.768**	-2.298***	-1.596***	-3.029**
Quality of roads (ln)	0.024	0.141	0.334***	0.384**	-0.885**	-0.693***	-0.662**	-0.685	0.443	0.084	0.060	0.082
Quality of port infrastructure (ln)	-0.426***	-0.191*	0.046	0.172	1.479***	1.481***	1.230**	0.659	0.167	1.117***	1.309***	2.528***
Quality of air infrastructure (ln)	-0.062	-0.213*	-0.347**	-0.626***	-1.196**	-1.424***	-1.177***	0.651	-0.317	-0.447	-0.613*	-1.814**
Number of documents to export	0.018	0.022	0.000	0.014	0.0676	0.157**	0.196***	0.194	-0.0298	0.053**	0.060*	0.140
Cost to export (US$ per container) (ln)	0.044	-0.038	-0.180	-0.208	1.065	0.352	-0.0479	0.053	0.247	0.242	0.490	3.509*
Number of documents to import	-0.002	-0.004	-0.022	-0.028	-0.0884	-0.054	-0.0751	-0.140**	0.0126	0.018	0.009	-0.006
Cost to import (US$ per container) (ln)	0.100	0.045	0.088	0.130	-0.626	-0.396	-0.13	-0.142	-0.58	-0.093	-0.218	-2.671*
WB income group of js partners	-	0.230***	0.233***	0.233***	-	0.241***	0.240***	0.241***	-	0.279***	0.279***	0.264***
Constant	-0.381	1.819***	3.582***	3.726***	-1.496	2.559**	3.431***	2.229	6.053*	2.119	0.842	-1.783
Individual fixed effects	YES	YES	NO	NO	YES	YES	NO	NO	YES	YES	NO	NO
Time fixed effects	YES	YES	YES	NO	YES	YES	YES	NO	YES	YES	YES	NO
Observations	828	828	828	248	218	202	202	56	224	224	224	64
Groups	248	248			56	52			64	64		
F-statistic overall significance	3.19	675.12	49.83	57.35	1.31	524.86	28.50	32.59	1.49	1039.92	67.66	90.84
F-statistic overall significance (P-value)	0.000	0.000	0.000	0.000	0.236	0.000	0.000	0.000	0.145	0.000	0.000	0.000
F-test on fixed effects (P-value)	0.000				0.000				0.000			
Breusch-Pagan LM test on random effects (P-value)		0.000				0.000				0.000		
Hausman test (FE vs RE)		0.000				0.027				0.007		
Sargan-Hansen Statistic (FE vs RE)		0.000				0.000				0.001		
R-squared	0.897	0.624	0.658	0.739	0.877	0.752	0.761	0.873	0.929	0.856	0.863	0.925

Note: Asterisks denote significance levels; * p<0.10, ** p<0.05 and *** p<0.01. P-values based on clustered standard errors

Table 4.17 Estimation results with quality of overall infrastructure and Trading Across Border score

Dependent variable: Average trade costs (ln), σ=8	Landlocked				Coastal				Partial-insularity				Island-states			
	FE	RE	Pooled	Cross-section	FE	RE	Pooled	Cross-section	FE	RE	Pooled	Cross-section	FE	RE	Pooled	Cross-section
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Ruggedness (ln)		0.081**	0.047	0.051		0.009	0.014	0.030		0.013	0.089*	0.145**		0.021	0.044	-0.001
Average temperature (ln)		0.089	0.078	0.110		0.202***	0.181***	0.132**		0.176**	0.323***	0.309**		0.052	0.056	0.023
Distance from Equator (ln)		0.006	0.041	0.102**		-0.087***	-0.063**	-0.024		0.017	0.053	0.047		-0.021	-0.053	-0.022
Logistic Performance Index (ln)	-0.413***	-1.129***	-1.977***	-2.817***	-0.228	-0.751***	-1.339***	-1.599***	-0.119	-0.212	-0.832	-2.224*	0.405**	-0.007	-0.381	-0.268
Linear Shipping Connectivity Index (ln)					-0.027	-0.219***	-0.249***	-0.284***	0.041	-0.318***	-0.428***	-0.465***	-0.133	-0.197***	-0.202***	-0.180***
Global Competitiveness Index (ln)	1.178***	-0.51	-0.391	-0.742	1.488***	0.279	0.159	0.085	0.196	-0.68	0.384	1.187	-1.554**	-3.273***	-2.977***	-3.751***
Quality of overall infrastructure (ln)	-0.261*	-0.305**	-0.218	0.298	-0.407**	-0.178	0.146	0.246*	0.173	0.006	-0.401	0.169	0.358	0.881***	1.103***	1.444***
Trading Across Border score (ln)	0.009	-0.077**	-0.112***	-0.075	-0.150	-0.116	-0.070	-0.235**	0.202	-0.361	0.247	0.251	-0.014	-0.105	-0.212	-0.418
WB income group of js partners		0.205***	0.211***	0.200***		0.224***	0.231***	0.229***		0.241***	0.240***	0.241***		0.266***	0.273***	0.252***
Constant	1.275***	3.957***	4.298***	5.063***	1.052	2.361***	2.436***	3.819***	-0.063	4.382***	1.437	1.137	3.423***	5.396***	5.487***	7.001***
Individual fixed effects	YES	YES	NO	NO	YES	YES	NO	NO	YES	YES	NO	NO	YES	YES	NO	NO
Time fixed effects	YES	YES	YES	NO	YES	YES	YES	NO	YES	YES	YES	NO	YES	YES	YES	NO
Observations	372	356	356	104	864	860	860	280	218	202	202	56	232	232	232	72
Groups	108	104			284	280			56	52			72	72		
F-statistic overall significance	5.56	314.76	56.05	68.69	4.19	734.61	72.02	117.48	0.42	213.29	28.92	38.4	1.63	941.34	57.27	71.79
F-statistic overall significance (P-value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.901	0.000	0.000	0.000	0.132	0.000	0.000	0.000
F-test on fixed effects (P-value)	0.000				0.000				0.000				0.000			
Breusch-Pagan LM test on random effects (P-value)		0.000				0.000				0.000				0.000		
Hausman test (FE vs RE)		0.000				0.000				0.000				0.000		
Sargan-Hansen Statistic (FE vs RE)		0.000				0.000				0.000				0.000		
R-squared	0.954	0.742	0.765	0.824	0.899	0.624	0.641	0.729	0.870	0.697	0.715	0.830	0.928	0.830	0.840	0.903

Note: Asterisks denote significance levels; * $p<0.10$, ** $p<0.05$ and *** $p<0.01$. P-values based on clustered standard errors

Table 4.18 Estimation results: cross-section for single years (2007 and 2014) with trade costs computed using σ=7 and σ=11

	Landlocked				Coastal				Partial-insularity				Island-states			
	2007	2007	2014	2014	2007	2007	2014	2014	2007	2007	2014	2014	2007	2007	2014	2014
Dependent variable: Average trade costs (ln)	σ=7	σ=11	σ=7	σ=11	σ=7	σ=11	σ=7	σ=11	σ=7	σ=11	σ=7	σ=11	σ=7	σ=11	σ=7	σ=11
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Ruggedness (ln)	0.092	0.078*	0.047	0.039	-0.048	-0.034	0.067**	0.043**	0.156	0.135	0.339**	0.222**	0.214**	0.149**	0.145	0.098
Average temperature (ln)	0.203*	0.155*	0.044	0.026	0.083	0.049	0.212***	0.122**	0.861	0.800	-0.329	-0.247	-0.240	-0.149	0.145	0.102
Distance from Equator (ln)	0.285**	0.220**	-0.043	-0.033	-0.090*	-0.069*	-0.078	-0.060	-0.059	-0.106	0.431***	0.289***	-0.599***	-0.385***	-0.219	-0.140
Logistic Performance Index (ln)	-0.919**	-0.615**	-2.688***	-1.942***	-1.333***	-0.910***	-2.805***	-1.982***	-0.171	0.985			-0.965***	-0.670***	-1.941**	-1.308**
Linear Shipping Connectivity Index (ln)					-0.161**	-0.121***	-0.229***	-0.166***	-0.876	-0.856	0.177	0.138	-0.438***	-0.307***	-0.247**	-0.177**
Global Competitiveness Index (ln)	0.434	0.263	-0.809	-0.717	-1.820***	-1.390***	1.445**	0.899**	-11.810	-12.180	0.840	0.659			-0.118	-0.055
Quality of overall infrastructure (ln)	-0.888***	-0.653***	0.909***	0.698***	0.466*	0.350**	-0.311	-0.220	5.162	5.817	0.506	0.172	1.231	0.972	0.828	0.516
Quality of air infrastructure (ln)	-0.058	-0.043	-0.662**	-0.478**	-0.038	-0.042	0.435	0.326	0.472	-0.361	-3.231	-2.192	-0.309	-0.392	-0.104	0.017
Number of documents to export	0.079**	0.058**	0.008	0.006	0.007	0.010	0.0852*	0.069**	-0.169	-0.143	0.145	0.107	-0.056	-0.060	0.116**	0.091**
Cost to export (US$ per container) (ln)	-0.583*	-0.463*	0.052	0.041	-0.230*	-0.137	-0.125	-0.090	3.744	3.888	-1.760	-1.134			1.52	1.022
Number of documents to import	-0.026	-0.020	0.053	0.038	-0.048**	-0.037**	-0.045	-0.042*	-0.117	-0.148	0.075	0.0471	0.093	0.086*	-0.035	-0.019
Cost to import (US$ per container) (ln)	0.865***	0.674***	0.084	0.049	-0.030	-0.042	-0.130	-0.070	-4.030	-3.899	0.643	0.412	0.602**	0.465**	-0.586	-0.391
WB income group of js partners	0.238***	0.185***	0.201***	0.160***	0.238***	0.187***	0.247***	0.198***	0.268***	0.199***	0.246***	0.197***	0.322***	0.243***	0.285***	0.217***
Constant	0.464	-0.325	3.823***	2.329***	7.029***	4.442***	4.008***	2.288***	14.490*	11.320*	11.55***	7.103***	-2.206	-2.357**	-3.972	-3.414
Observations	88	88	100	100	192	192	212	212	50	50	48	48	44	44	60	60
R-squared	0.858	0.853	0.829	0.827	0.713	0.729	0.641	0.671	0.758	0.785	0.863	0.881	0.892	0.902	0.908	0.914

Note: Asterisks denote significance levels; * $p<0.10$, ** $p<0.05$ and *** $p<0.01$. P-values based on clustered standard errors

CONCLUDING REMARKS

Trade costs and, in particular, the indirect approach to compute them, have driven the writing of this thesis.

Their importance for development and for economic performance of countries, the big role they play in explaining international trade patterns, and the need of further research are the three main reasons that have led to the choice of devoting this dissertation to the trade cost subject.

The contribution that this thesis has presented in terms of understanding and empirical results has not been plain. Indeed, the main theme of trade costs has been investigated in conjunction with the historical and geographical perspective. The significant function that history and geography perform in determining modern economic outcomes has represented the background of the entire discussion.

In order to provide, on the one hand, a complete and exhaustive overview of the trade costs subject and, on the other hand, new empirical evidence about why and what makes trade costly, this thesis has been structured in a theoretical survey, two empirical applications and one instrumental section that supports the first empirical investigation.

To introduce the reader to the vast trade costs topic, this thesis has opened the discussion with an extensive and general survey mainly linked with the theoretical issues about trade costs. In these terms, it has been possible to appreciate why trade costs are able to explain different and detailed aspects of international trade and of economics in general. Trade costs describe why some countries trade internationally and others don't, why across years trade has experienced booms and busts, why during last thirty years countries have become more integrated and trade across them has grown exponentially. The survey provided has highlighted how both ways to model trade costs, ad valorem and per unit, are consistent with the data and able to capture the structure of trade costs. Analogously, direct and indirect methods to compute trade costs are both reliable. The choice of the best approach, the most suitable in measuring trade costs, is not trivial and is strictly correlated with the data availability. Both methods entail advantages and disadvantages. The direct measurement is straightforward in small analysis, with few countries and where all the information about singles sub-components is accessible. However, non-neutral assumptions (i.e. independence between variables) are needed to perform a measurement based on the adding up of heterogeneous elements. The indirect approach is the most appropriate when dealing with a large sample of countries and with long-term analyses. However, it does not allow to distinguish the contribution of each single trade cost factor and some weaknesses arise when dealing with sub-national level investigations. A preliminary discussion on new issues linked to indirect method to compute trade costs has represented the novelty of the first chapter. These 'new' thoughts need further research and further discussion to mitigate the limitedness of the measure and to enhance the strength of the indirect measurement. Intended work will move in the direction of interregional trade costs. Exploiting available data on interregional trade across the 276 NUTS2 EU regions, the aim is to capture the extent of trade costs across European regions and to investigate the potential of the measure at the sub-national level.

A pioneering investigation on sub-national trade costs has been provided in the first empirical application of this thesis. It has aimed at studying the long-term effects of the old Roman road network on present economic outcomes, exploiting trade costs computed for the Italian provinces, according to the indirect approach proposed by Novy (2013), as modern economic measure. Differently from previous works, that used measures based mainly on development (like GDP or population), this has been the first contribution, in the literature on the persistency of history, that has employed trade costs at the provincial level to assess whether the Roman road infrastructure left a mark on current economy. In order to measure the phenomenon in the best possible way, a measure of Roman road in kilometres has been computed for each Italian province, providing an approach that can be exploited also for future research. Since the significance of the theme and of the new constructed data set, the first empirical application has been addressed in two chapters. One chapter has reviewed the literature on the heritage left by historical episodes on present economic development, providing also evidence of historical facts that did not leave a mark on current economy, exploring the Italian history and which historical events had an impact on modern Italian economic outcomes, and examining all those works that focused on the long-run effects of 'Romanness'. Moreover, a description of the new constructed data set and of the methodology adopted to create it has been presented. With this extensive background, a dedicated chapter has been devoted to the sole analysis of the impact of the Roman road system on Italian provincial trade costs. The interest on what has been the long-run effect of Roman roads has been investigated in both direct and indirect terms. The first issue has been to assess the pure and comprehensive effect of the Roman infrastructure, ignoring the channels through which Roman roads performed. In a second stage, the focus was mainly on the current infrastructure and the Roman road measure has been used as an instrumental variable for modern roads, in order to assess whether Roman roads performed mainly through their effect on infrastructure. The evidence that has emerged is strong. The Roman road network has led to current lower trade costs: Italian provinces with a denser Roman road system are more prone to trade more internationally than domestically. The main idea behind it is that provinces crossed by the Roman road system have benefited of the Roman domination not only in terms of infrastructure, but also in terms of a more ʽmental openness', due to developed urban centres, flourishing economic activities and intense trade relationships. These promising results will lead future research to further investigate the Roman road theme. The idea is to compute the Roman road measure for the 276 NUTS2 EU regions and to combine the main theme of the persistent effect of the Roman road network with interregional trade costs for the European regions.

The second empirical application has adopted the more traditional bilateral measure of trade costs at the country level to investigate what the main determinants of trade costs are. Consistent with the first empirical contribution, trade costs have been computed according to the ʽtop-down' approach proposed by Novy (2013). The interest on the main sources that make trade costly has been focused on the geography of countries, distinguishing them according to their ʽdegree of insularity'. The choice of constituting four geographical groups of countries has been driven by the highly documented importance that geography plays in international trade and in the economic development of nations. On these bases, the attention has been focused not only on the determinants of trade costs, but also on whether these determinants vary by geographies of countries. Differently from the first contribution, the second empirical chapter has involved more theoretical insights linked to the indirect measure of trade costs. First, a choice between the ʽadjusted' and the ʽunadjusted' version of trade costs has been made, exploiting the ʽunadjusted' version for the main analysis and the ʽadjusted' one only for robustness checks. Second, since the potential given by the combination of the Heckman model approach and the ʽunadjusted' version, it has been possible to investigate the case of zero trade flows. Both descriptive and inferential analyses have highlighted how, having a disadvantageous geographical condition reflects on both participation to international exchanges and trade costs. Landlocked countries and island-states face the major difficulties in terms of trading internationally and high trade costs. However, whereas landlocked countries are able to overcome the unfavourable geography when they are rich, this is less the case for countries completely surrounded by the ocean. Having an access to sea or having islands (but not being an island) are the two most advantageous geographical conditions. Bilateral distance, logistic abilities, competitiveness marks are the factors that matter most in global trade and that make trade costly for all geographical groups. Maritime connectivity, the cost to export and trade facilitation bottlenecks, although significant, have a more marginal role. For landlocked and coastal countries' trade costs the main source is represented by the logistic abilities. Overall performance and economic competitiveness really matter for island-states, where, surprisingly, maritime connectivity is not one of the most important determinants. This first important evidence is intended to be investigated with future research relying on Bayesian Model Averaging (BMA), to deal with the potential problem of model uncertainty that affects the analysis of the determinants of trade costs. Nevertheless, the investigation performed has underlined how trade costs analyses should consider the geographical perspective to better appreciate the pattern of trade costs.

In terms of policy recommendations, this thesis highlights how governments should take history into account and how important are ad hoc policies promptly implemented after the historical event takes places. These policies should be designed in order to mitigate the negative effects of history and reduce the persistence that many historical facts entail. If it is not possible to go back to the past, it is possible to learn from the past. Investments in infrastructure and integration measures are fundamental in making trade dynamic and trade costs lower. Moreover, countries should invest in logistics and in all those abilities that enhance the performance of countries and that connect them with the rest of the world. If physical distances and geographical features cannot be improved or cannot be cancelled out, focused measures in maritime connectivity and in logistic services can not only alleviate the weaknesses coming from a disadvantageous geographical condition, but also boost competitiveness of countries.

BIBLIOGRAPHY

[1] Acemoglu, D., S. Johnson and J.A. Robinson (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation", *The American Economic Review*, 91(5), pp. 1369-1401.

[2] Acemoglu, D., S. Johnson and J.A. Robinson (2002), "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution", *The Quarterly Journal of Economics*, 117(4), pp. 1231-1294.

[3] Afonso, O. (2001), "The Impact of International Trade on Economic Growth", Investigação–Trabalhos em Curso, 106.

[4] Alchian, A.A. and W.R. Allen (1964), *University Economics*, Belmont, California: Wadsworth.

[5] Anderson, J.E. and D. Marcouiller (2002), "Insecurity and the Pattern of Trade: An Empirical Investigation", *The Review of Economics and Statistics*, 84(2), pp. 342-352.

[6] Anderson, J.E. and E. van Wincoop (2002), "Borders, Trade, and Welfare", in: Collins, S.M., Rodrik, D. (Eds.), Brookings Trade Forum, 2001, Washington DC: Brookings Institution, pp. 207-244.

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

[8] Anderson, J.E. and E. van Wincoop (2004), "Trade Costs", *Journal of Economic Literature*, 42(3), pp. 691-751.

[9] Anderson, J.E. and Y. Yotov (2010), "The Changing Incidence of Geography", *The American Economic Review*, 100(5), pp. 2157-2186.

[10] Anderson, J.E. and Y. Yotov (2012), “Gold Standard Gravity”, NBER Working Paper, w17835.

[11] Arabmazar, A. and P. Schmidt (1981), “Further Evidence on the Robustness of the Tobit Estimator to Heteroskedasticity”, *Journal of Econometrics*, 17(2), pp. 253-258.

[12] Armington, P.S. (1969), “A Theory of Demand for Products Distinguished by Place of Production”, *Staff Papers*, 16(1), pp. 159-178.

[13] Armstrong, H., R.J. De Kervenoael, X. Li and R. Read (1998), “A Comparison of the Economic Performance of Different Micro-States, and between Micro-States and Larger Countries”, *World Development*, 26(4), pp. 639-656.

[14] Arvis, J.F., Y. Duval, B. Shepherd and C. Utoktham (2013a), “Trade Costs in Developing World: 1995-2010”, World Bank Policy Research Working Paper 6309.

[15] Arvis, J.F., Y. Duval, B. Shepherd and C. Utoktham (2013b), “Trade Costs and Development: A New Data Set”, Economic Premise No. 104, Washington, DC: World Bank.

[16] Arvis, J.F., D. Saslavsky, L. Ojala, B. Shepherd, C. Busch and A. Raj (2014), *Connecting to Compete 2014: Trade Logistics in the Global Economy*, Washington, DC: World Bank.

[17] Arvis, J.F., G. Raballand and J.F. Marteau (2010), *The Cost of Being Landlocked: Logistics Costs and Supply Chain Reliability*, Washington, DC: World Bank.

[18] Aschauer, D.A. (1990), “Why Is Infrastructure Important?”, *Conference Series*, Federal Reserve Bank of Boston, 34, pp. 21-68.

[19] Atack, J., F. Bateman, M. Haines and R.A. Margo (2010), “Did Railroads Induce or Follow Economic Growth? Urbanization and Population Growth in the American Midwest, 1850-1860”, *Social Science History*, 34(2), pp. 171-197.

[20] Automobile Club Italia (2011), “Dotazione di Infrastrutture Stradali sul Territorio Italiano”, available at (accessed 25 July 2016): http://www.aci.it/fileadmin/documenti/studi_e_ricerche/dati_statistiche/Infrastrutture_stradali_in_Italia/Dotazione_di_infrastrutture_stradali_in_Italia.pdf.

[21] Aw, W.B. (1981), *Highway Construction Cost Model for Sector Planning in Developing Countries*, S.M. thesis, Cambridge, MA, Massachusetts Institute of Technology.

[22] Baier, S.L. and J.H. Bergstrand (2001), "The Growth of World Trade: Tariffs, Transport Costs and Income Similarity", *Journal of International Economics*, 53(1), pp. 1-27.

[23] Baldwin, R.E. and V. DiNino (2005), "Euros and Zeros: The Common Currency Effect on Trade in New Goods", CEPR Discussion Papers 5973.

[24] Baldwin, R.E., R. Forslid, P. Martin and F. Robert-Nicoud (2003), "The Core-Periphery Model: Key Features and Effects", in: Heijdra and Brakman, Eds., *The Monopolistic Competition Revolution in Retrospect*, Cambridge University Press, pp. 213-235.

[25] Baldwin, R.E. and P. Martin (1999), "Two Waves of Globalisation: Superficial Similarities, Fundamental Differences", (No. w6904), National Bureau of Economic Research.

[26] Banerjee, A., E. Duflo and N. Qian (2012), "On the Road: Access to Transportation Infrastructure and Economic. Growth in China", NBER Working Paper N. 17897.

[27] Baum, C.F., M.E. Schaffer and S. Stillman (2003), "Instrumental Variables and GMM: Estimation and Testing", *The Stata Journal*, 3(1), pp. 1-31.

[28] Becker, C. (2012), "Small Island States in the Pacific: The Tyranny of Distance", IMF Working Paper, 223.

[29] Beghin, J. (2008), "Nontariff Barriers", in The New Palgrave Dictionary of Economics, ed. S. Darlauf and L. Blume, 2nd ed., New York: Palgrave Macmillan pp. 126-129.

[30] Behar, A. and A.J. Venables (2010), "Transport Costs and International Trade", Economics Series Working Papers 488, University of Oxford, Department of Economics.

[31] Berger, T. and K. Enflo (2015), "Locomotives of Local Growth: The Short- and Long-Term Impact of Railroads in Sweden", *Journal of Urban Economics* (2015).

[32] Bertram, G. and O. Karagedikli (2004), *Are Pacific Economies Converging or Diverging?*, In Poot, J. (ed), On the Edge of the Global Economy, Cheltenham: Edward Elgar, pp. 106-122.

[33] Bishop, M.C. (2014), *The Secret History of the Roman Roads of Britain*, Pen & Sword Military.

[34] Bleaney, M. and A. Dimico (2008), "Geography Matters: Reconsidering the Effect of Geography on Development", *Discussion Papers* 08/14, University of Nottingham, CREDIT.

[35] Bloom, D.E. and J.G. Williamson (1998), "Demographic Transitions and Economic Miracles in Emerging Asia", *The World Bank Economic Review*, 12(3), pp. 419-455.

[36] Borgatti, L. (2008), "Pacific Islands' Bilateral Trade: The Role of Remoteness and of Transport Costs", *Journal of International Development*, 20(4), pp. 486-501.

[37] Bosker, M. and E. Buringh (2012), "City Seeds: Geography and the Origins of the European City System", Mimeo.

[38] Bosker, M., E. Buringh and J.L. van Zanden. (2013), "From Baghdad to London: Unraveling Urban Development in Europe, the Middle East, and North Africa, 800-1800", *The Review of Economics and Statistics*, 95(4), pp. 1418-1437.

[39] Bouet, A., S. Mishra and D. Roy (2008), "Does Africa Trade Less Than It Should, and If So, Why?: The Role of Market Access and Domestic Factors", International Food Policy Research Institute Discussion Paper No. 770.

[40] Bowsher, C.G. (2002), "On Testing Overidentifying Restrictions in Dynamic Panel Data Models", *Economics Letters*, 77(2), pp. 211-220.

[41] Briguglio, L. (1995), "Small Island Developing States and their Economic Vulnerabilities", *World Development*, 23(9), pp. 1615-1632.

[42] Briguglio, L. and J. Kaminarides (1993), *Islands and Small States: Issues and Policies*, in a special issue of World Development. New York: Pergamon Press.

[43] Broda, C. and D.E. Weinstein (2006), "Globalization and the Gains from Variety", *The Quarterly Journal of Economics*, 121(2), pp. 541-585.

[44] Brooks, D.H. and D. Hummels (2009), *Infrastructure's Role in Lowering Asia's Trade Costs: Building for Trade*, Edward Elgar Publishing.

[45] Buringh, E., J.L van Zanden and M. Bosker (2012), "Soldiers and Booze: The Rise and Decline of a Roman Market Economy in North-Western Europe", Utrecht University, CGEH Working Paper No. 32.

[46] Cameron, A.C. and P.K. Trivedi (2005), *Microeconometrics: Methods and Applications*, New York: Cambridge University Press.

[47] Cantoni, D. (2015), "The Economic Effects of the Protestant Reformation: Testing the Weber Hypothesis in the German Lands", *Journal of the European Economic Association*, 13(4), pp. 561-598.

[48] Chambers, W. (1860), *Chambers's Encyclopædia*, London: George Newnes, (1961).

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

[50] Chatfield, C. (1995), "Model Uncertainty, Data Mining, and Statistical Inference", *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 158(3), pp. 419-466.

[51] Chen, N.A. and D. Novy (2011), "Gravity, Trade Integration, and Heterogeneity across Industries", *Journal of International Economics*, 85(2), pp. 206-221.

[52] Chen, N.A. and D. Novy (2012), "On the Measurement of Trade Costs: Direct vs. Indirect Approaches to Quantifying Standards and Technical Regulations", *World Trade Review*, Cambridge University Press, 11(3), pp. 401-414.

[53] Chevallier, R. (1976), *Roman Roads*, trans. N.H. Field, Berkeley: University of California press, 1976.

[54] Cole, M.T. (2011), "Not All Trade Restrictions are Created Equally", *Review of World Economics*, Springer Verlag, 147(3), pp. 411-427.

[55] Clemente, G. (1990), *Basi Sociali e Assetti Istituzionali nell'Età della Conquista*, in Arnaldo Momigliano; Aldo Schiavone (edited by), Storia di Roma Vol. II/1, Torino, Einaudi.

[56] Cornell, T. (1995), *The Beginnings of Rome: Italy and Rome from the Bronze Age to the Punic Wars*, Routledge.

[57] Cornell, T. and J. Matthews (1982), *Atlas of the Roman World*, New York, N.Y.: Facts on File.

[58] Cragg, J.G. and S.G. Donald (1993), "Testing Identifiability and Specification in Instrumental Variable Models", *Econometric Theory*, 9(2), pp. 222-240.

[59] Crespo Cuaresma, J., G. Doppelhofer, and M. Feldkircher (2009), "The Determinants of Economic Growth in European Regions", CESifo Working Paper Series No. 2519.

[60] Crosby, A.W. (1986), *Ecological Imperialism: The Biological Expansion of Europe, 900-1900*, Cambridge, UK: Cambridge University Press.

[61] Dalgaard, C.J., N. Kaarsen, O. Olsson and P. Selaya (2015), "Roman Roads to Prosperity: The Long-run Impact of Transport Infrastructure", SLIDE PRESENTATION available at (accessed 01 December 2015): http://economics.gu.se/digitalAssets/1531/1531740_roman4.pdf.

[62] Davies, H.E.H. (1998), "Designing Roman Roads", *Britannia*, 29, pp. 1-16.

[63] Davis, D.R and D.E Weinstein (2002), "Bones, Bombs, and Break Points: The Geography of Economic Activity", *The American Economic Review*, 92(5), pp. 1269-1289.

[64] De Benedictis, L. and R. Helg (2002), "Globalizzazone", *Rivista di Politica Economica*, 92(3/4), pp. 139-210.

[65] De Benedictis, L. and A.M. Pinna (2015), "Islands as "Bad Geography". Insularity, Connectedness, Trade Costs and Trade", Working Paper Crenos 2015/04.

[66] De Benedictis, L. and D. Taglioni (2011), "The Gravity Model in International Trade", in *The Trade Impact of European Union Preferential Policies: An Analysis through Gravity Models*, Luca De Benedictis and Luca Salvatici (eds.), Springer.

[67] De Francisci, P. (1968), *Sintesi Storica del Diritto Romano*, Bulzoni.

[68] De la Roca, J. and D. Puga (2017), "Learning by Working in Big Cities", *The Review of Economic Studies*, 84(1), pp. 106-142.

[69] Deardorff, A.V. (2014), "Local Comparative Advantage: Trade costs and the Pattern of Trade", *International Journal of Economic Theory*, 10(1), pp. 9-35.

[70] Deardorff, A.V. and R.M. Stern (1998), *Measurement of Nontariff Barriers*, Ann Arbor: University of Michigan Press.

[71] Del Bo, C.F. and M. Florio (2012), "Infrastructure and Growth in a Spatial Framework: Evidence from the EU Regions", *European Planning Studies*, 20(8), pp. 1393-1414.

[72] Dell, M. (2010), "The Persistent Effects of Peru's Mining Mita", *Econometrica*, 78(6), pp. 1863-1903.

[73] Dell, M., B.F. Jones and B.A. Olken (2012), "Temperature Shocks and Economic Growth: Evidence from the Last Half Century", *American Economic Journal: Microeconomics*, 4(3), pp. 66-95.

[74] Di Liberto, A. and M. Sideri (2015), "Past Dominations, Current Institutions and the Italian Regional Economic Performance", *European Journal of Political Economy*, Elsevier, 38(C), pp. 12-41.

[75] Diamond, J. (1997), *Guns, Germs and Steel: The Fates of Human Societies*, Norton.

[76] Dijkstra, T.K. (1988), *On Model Uncertainty and Its Statistical Implications*, Springer, Berlin.

[77] Disdier, A. and S. Marette (2010), "The Combination of Gravity and Welfare Approaches for Evaluating Nontariff Measures", *American Journal of Agricultural Economics*, 92(3), pp. 713-726.

[78] Dixit, A.K. and J.E. Stiglitz (1977), "Monopolistic Competition and Optimum Product Diversity", *The American Economic Review*, 67(3), pp. 297-308.

[79] Doepke, M. and F. Zilibotti (2005), "Social Class and the Spirit of Capitalism", *Journal of the European Economic Association*, 3(2-3), pp. 516-524.

[80] Doepke, M. and F. Zilibotti (2008), "Occupational Choice and the Spirit of Capitalism", *The Quarterly Journal of Economics*, 123(2), pp. 747–793.

[81] Dommen, E. and P. Hein (1985), *States, Microstates and Islands*, Croom Helm, London, U.K.

[82] Donaldson, D. (2015), "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure", *The American Economic Review*, forthcoming.

[83] Donaldson, D. and R. Hornbeck (2016), "Railroads and American Economic Growth: A 'Market Access' Approach", *The Quarterly Journal of Economics*, 131(2), pp. 799-858.

[84] Dornbusch, R., S. Fischer and P.A. Samuelson (1977), "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods", *The American Economic Review*, 67(5), pp. 823-839.

[85] Draper, D. (1995), "Assessment and Propagation of Model Uncertainty", *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 45-97.

[86] Draper, D., J.S. Hodges, E.E. Leamer, C.N. Morris and D.B. Rubin (1987), "A Research Agenda for Assessment and Propagation of Model Uncertainty", Technical Report Rand Note N-2683-RC, RAND Corporation, Santa Monica, California.

[87] Duval, Y. and C. Utoktham (2011a), "Intraregional Trade Costs in Asia: A Primer", *Asia-Pacific Development Journal*, 18(2), pp. 1-23.

[88] Duval, Y. and C. Utoktham (2011b), "Trade Facilitation in Asia and the Pacific: Which Policies and Measures Affect Trade Costs the Most?", Asia-Pacific Research and Training Network on Trade Working Paper Series, No. 94.

[89] Easterly, W. and R. Levine (2003), "Tropics, Germs, and Crops: How Endowments Influence Economic Development", *Journal of Monetary Economics*, 50(1), pp. 3-39.

[90] Eaton, J. and S. Kortum (2002), "Technology, Geography, and Trade", *Econometrica*, 70(5), pp. 1741-1779.

[91] Emlinger, C., F. Jacquet and E.C. Lozza (2008), "Tariffs and Other Trade Costs: Assessing Obstacles to Mediterranean Countries' Access to EU-15 Fruit and Vegetable Markets", *European Review of Agricultural Economics*, 35(4), pp. 409-438.

[92] Engerman, S.L. and K.L. Sokoloff (1997), "Factor Endowments, Institutions, and Differential Paths of Growth among New World Economies: A View from Economic Historians of the United States", in How Latin America Fell Behind, ed. S. Harber, Stanford, CA: Stanford University Press, pp. 260-304.

[93] Engerman, S.L. and K.L. Sokoloff (2000), "Institutions, Factor Endowments, and Paths of Development in the New World", *The Journal of Economic Perspectives*, 14(3), pp. 217-232.

[94] Engerman, S.L. and K.L. Sokoloff (2002), "Factor Endowments, Inequality, and Paths of Development Among New World Economies", *Economia,* 3(1), pp. 41-110.

[95] Estervadeordal, A., B. Frantz and A.M. Taylor (2003), "The Rise and Fall of World Trade, 1870-1939", *The Quarterly Journal of Economics*, 118(2), pp. 359-407.

[96] Faber, B. (2014), "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System", *The Review of Economic Studies*, 81(3), pp. 1046-1070.

[97] Fajgelbaum, P. and S.J. Redding (2014), "External Integration, Structural Transformation and Economic Development: Evidence from Argentina 1870-1914", NBER Working Paper No. 20217.

[98] Fally, T. (2015), "Structural Gravity and Fixed Effects", *Journal of International Economics*, 97(1), pp. 76-85.

[99] Faye, M.L., J.W McArthur, J.D. Sachs and T. Snow (2004), "The Challenges Facing Landlocked Developing Countries"*, Journal of Human Development*, 5(1), pp. 31-68.

[100] Feenstra, R.C. (1994), "New Product Varieties and the Measurement of International Prices", *The American Economic Review*, 84(1), pp. 157-177.

[101] Feenstra, R.C. and J. Romalis (2014), "International Prices and Endogenous Quality", *The Quarterly Journal of Economics*, 129(2), pp. 477-527.

[102] Fernández, C., E. Ley and M.F. Steel (1997), "Statistical Modeling of Fishing Activities in the North Atlantic", No. 97-25, Fedea.

[103] Fernández, C., E. Ley and M.F. Steel (2001), "Benchmark Priors for Bayesian Model Averaging", *Journal of Econometrics*, 100(2), pp. 381–427.

[104] Finley, M.I. (1973), *The Ancient Economy*, Berkeley and Los Angeles: University of California Press.

[105] Fogel, R.W. (1964), *Railroads and American Economic Growth: Essays in Econometric History*, Baltimore: Johns Hopkins University Press.

[106] Fontana, N., T. Nannicini and G. Tabellini (2016), "Historical Roots of Political Extremism: The Effects of Nazi Occupation of Italy", Working Paper.

[107] Fosbroke, T.D. (1843), *Encyclopaedia of Antiquities, and Elements of Archaeology, Classical and Mediaeval Volume 2*, M.A. Nattali, London.

[108] Fouquin, M. and J. Hugot (2014), "Trade Globalization: 1827-2012. When Did Trade Costs Start to Fall?", CEPII Working Papers.

[109] Fouquin, M. and J. Hugot (2016), "Back to the Future: International Trade Costs and the Two Globalizations", CEPII Working Papers 2016-13.

[110] Frankel, A. (1997), *Regional Trading Blocs in the World Economic System*, Washington DC: Peterson Institute for International Economics.

[111] Frieden, J.A. (2007), *Global Capitalism: Its Fall and Rise in the Twentieth Century*, W.W. Norton & Co.

[112] Fujita, M., P.R. Krugman and A.J. Venables (1999), "The Spatial Economy – Cities, Regions, and International Trade", Cambridge, MA: MIT Press.

[113] Gallup, J.L., J.D. Sachs and A.D. Mellinger (1999), "Geography and Economic Development", *International Regional Science Review*, 22(2), pp. 179-232.

[114] Garcia-López, M.À., A. Holl and E. Viladecans-Marsal (2015), "Suburbanization and Highways in Spain when the Romans and the Bourbons still Shape its Cities, *Journal of Urban Economics*, 85, pp. 52-67.

[115] Gaurav, A. and K. Mathur (2016), "Determinants of Trade Costs and Trade Growth Accounting between India and the European Union during 1995-2010", *The World Economy*, 39(9), pp. 1399-1413.

[116] Gedacht, D. (2004), *Land and Resources of Ancient Rome*, Rosen Publishing.

[117] Gleason, J.P. (2013), *Roman Roads in Gaul: How Lines of Communication and Basing Support Operational Reach*, School of Advanced Military Studies Monographs.

[118] Greene, W.H. (2008), *Econometric Analysis*, 6th ed., Upper Saddle River, NJ: Prentice Hall.

[119] Guiso, L., P. Sapienza and L. Zingales (2016), "Long-Term Persistence", *Journal of the European Economic Association*, 14(6), pp. 1401-1436.

[120] Hacquard, G. (2003), *Guia De La Roma Antigua*, Centro de Linguistica Aplicada Atenea.

[121] Hall, R.E. and C.I. Jones (1999), "Why Do Some Countries Produce So Much More Output per Workers than Others?", *The Quarterly Journal of Economics*, 114(1), pp. 83-116.

[122] Hallak, J.C. (2006), "Product Quality and the Direction of Trade", *Journal of International Economics*, 68(1), pp. 238-265.

[123] Haq, Z., K.D. Meilke and J. Cranfield (2011), "The Gravity Model and the Problem of Zero's in Agrifood Trade", CATPRN Working Paper 2011-02, Canadian Agricultural Trade Policy and Competitiveness Research Network, Guelph.

[124] Harley, C.K. (1988), "Ocean Freight Rates and Productivity, 1740-1913: The Primacy of Mechanical Invention Reaffirmed", *The Journal of Economic History*, 48(4), pp. 851-876.

[125] Harrigan, J. (1993), "OECD Imports and Trade Barriers in 1983", *Journal of International Economics*, 35(1-2), pp. 91-111.

[126] Hausman, J.A. (1978), "Specification Tests in Econometrics", *Econometrica*, 46(6), pp. 1251-1271.

[127] Head, K. and T. Mayer (2004), *The Empirics of Agglomeration and Trade*, Handbook of Regional and Urban Economics Vol. 4, Chapter 59, North-Holland.

[128] Head, K. and J. Ries (2001), "Increasing Returns vs National Product Differentiation as an Explanation for the Pattern of US-Canada Trade", *The American Economic Review*, 91(4), pp. 858-876.

[129] Heckman, J.J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, 47(1), pp. 153-161.

[130] Helpman, E., M. Melitz and Y. Rubinstein (2008), "Estimating Trade Flows: Trading Partners and Trading Volumes", *The Quarterly Journal of Economics*, 123(2), pp. 441-487.

[131] Hindle, P. (1998), *Medieval Roads and Tracks*, ed.3, Princes Risborough.

[132] Hingley, R. (2015), "Post-Colonial and Global Rome: the Genealogy of Empire" in Globalisation and the Roman World: World History, Connectivity and Material Culture, ed. Martin Pitts and Miguel John Versluys, New York: Cambridge University Press, pp. 32-46.

[133] Hodges, J.S. (1987), "Uncertainty, Policy Analysis, and Statistics", *Statistical Science*, 2(3), pp. 259-275.

[134] Hoekman, B. and A. Nicita (2011), "Trade Policy, Trade Costs, and Developing Country Trade", *World Development*, 39(12), pp. 2069-2079.

[135] Hoeting, J.A., D. Madigan, A.E. Raftery and C.T. Volinsky (1999), "Bayesian Model Averaging: A Tutorial", *Statistical Science*, 14(4), pp. 382-417.

[136] Hoffmann, R. and L. Kassouf (2005), "Deriving Conditional and Unconditional Marginal Effects in Log Earnings Equations Estimated by Heckman's Procedure", *Applied Economics*, 37(11), pp. 1303-1311.

[137] Huberman, M., C.M. Meissner and K. Oosterlinck (2015), "Technology and Geography in the Second Industrial Revolution: New Evidence from the Margins of Trade", *NBER Working Papers 20851*, National Bureau of Economic Research, Inc.

[138] Hummels, D. (2001), "Toward a Geography of Trade Costs", mimeo, Purdue.

[139] Hummels, D. and P.J. Klenow (2005), "The Variety and Quality of a Nation's Exports", *The American Economic Review*, 95(3), pp. 704-723.

[140] Hummels, D. and A. Skiba (2004), "Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture", *Journal of Political Economy*, 112(6), pp.1384-1402.

[141] Hurd, M. (1979), "Estimation in Truncated Samples When There Is Heteroscedasticity", *Journal of Econometrics*, 11(2-3), pp. 247-258.

[142] Imbs, J. and I. Méjean (2009), "Elasticity Optimism", CEPR Discussion Paper 7177.

[143] Irarrazabal, A., A. Moxnes and L.D. Opromolla (2015), "The Tip of the Iceberg: A Quantitative Framework for Estimating Trade Costs", *The Review of Economics and Statistics*, 97(4), pp. 777-792.

[144] Jacks, D.S. (2006), "What Drove 19th Century Commodity Market Integration?", *Explorations in Economic History*, 43(3), pp. 383-412.

[145] Jacks, D.S., C.M. Meissner and D. Novy (2008), "Trade Costs, 1870-2000", *The American Economic Review*, 98(2), pp. 529-534.

[146] Jacks, D.S., C.M. Meissner and D. Novy (2010), "Trade Costs in the First Wave of Globalization", *Explorations in Economic History*, 47(2), pp. 127-141.

[147] Jacks, D.S., C.M. Meissner and D. Novy (2011), "Trade Booms, Trade Busts, and Trade Costs", *Journal of International Economics*, 83(2), pp. 185-201.

[148] James, H. (2001), *The End of Globalization - Lessons from the Great Depression*, Harvard University Press.

[149] Jayasinghe, S., J.C. Beghin and G. Moschini (2010), "Determinants of World Demand for U.S. Corn Seeds: The Role of Trade Costs", *American Journal of Agricultural Economics*, 92(4), pp. 999-1010.

[150] Jedwab, R., E. Kerby and A. Moradi (2015), "History, Path Dependence and Development: Evidence from Colonial Railroads, Settlers and Cities in Kenya", *The Economic Journal*, forthcoming.

[151] Jennings, J. (2011), *Globalizations and the Ancient World*, Cambridge University Press.

[152] Johnson, A. (2008), "The Three Waves of Globalisation", *Journal of Nordregio*, 8, pp. 8-9.

[153] Jones, E.L. (1981), *The European Miracle: Environments, Economies, and Geopolitics in the History of Europe and Asia*, Cambridge, MA: Cambridge University Press.

[154] Kamarck, A.M. (1976), *The Tropics and Economic Development*, Baltimore and London: Johns Hopkins University Press.

[155] Kareem, F.O. (2013), "Modeling and Estimation of Gravity Equation in the Presence of Zero Trade: A Validation of Hypotheses Using Africa's Trade Data", EAAE Seminar Paper.

[156] Kass, R.E. and A.E. Raftery (1995), "Bayes Factors", *Journal of the American Statistical Association*, 90(430), pp. 773-795.

[157] Kessides, C. (1993), "The Contributions of Infrastructure to Economic Development: A Review of Experience and Policy Implications", World Bank discussion papers, No. WDP 213, Washington, DC: The World Bank.

[158] Knack, S. and P. Keefer (1995), "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures", *Economics and Politics*, 7(3), pp. 207-227.

[159] Knapton, J. (1996), "The Romans and Their Roads - The Original Small Element Pavement Technologists", in 5th International Concrete Block Paving Conference, Tel-Aviv, Israel.

[160] Korinek, J. (2008), "Clarifying Trade Costs in Maritime Transport", OECD Working Paper TAD/TC/WP (2008) 10, Paris.

[161] Krugman, P.R. (1980), "Scale Economies, Product Differentiation, and the Pattern of Trade", *The American Economic Review*, 70(5), pp. 950-959.

[162] La Porta, R., F. Lopez-de-Silanes and A. Shleifer (2008), "The Economic Consequences of Legal Origins", *Journal of Economic Literature*, 46(2), pp. 285-332.

[163] La Porta, R., F. Lopez-de-Silanes, A. Shleifer and R.W. Vishny (1997), "Legal Determinants of External Finance", *Journal of Finance*, 52(3), pp. 1131-1150.

[164] La Porta, R., F. Lopez-de-Silanes, A. Shleifer and R.W. Vishny (1998), "Law and Finance", *Journal of Political Economy*, 106(6), pp. 1113-1155.

[165] La Porta, R., F. Lopez-de-Silanes, A. Shleifer and R.W. Vishny (1999), "The Quality of Government", *Journal of Law, Economics, and Organization*, 15, pp. 222-279.

[166] Laurence, R. (1999), *The Roads of Roman Italy: Mobility and Cultural Change*, Routledge, London.

[167] Laurence, R. and F. Trifilò (2015), "The Global and the Local in the Roman Empire: Connectivity and Mobility from an Urban Perspective" in Globalisation and the Roman World: World History, Connectivity and Material Culture, ed. Martin Pitts and Miguel John Versluys, New York: Cambridge University Press, pp. 99-122.

[168] Leamer, E.E. (1978), *Specification Searches: Ad Hoc Inference with Nonexperimental Data*, Wiley, New York.

[169] Leung, S.F. and S. Yu (1996), "On the Choice Between Sample Selection and Two-Part Models", *Journal of Econometrics*, 72(1), pp. 197-229.

[170] Limão, N. and A.J. Venables (2001), "Infrastructure, Geographical Disadvantage, Transport Costs, and Trade", *The World Bank Economic Review*, 15(3), pp. 451-479.

[171] Linders, G.M and H. de Groot (2006), "Estimation of the Gravity Equation in the Presence Of Zero Flows", Tinbergen Institute Discussion Papers 06-072/3.

[172] Linnemann, H. (1966), *An Econometric Study of International Trade Flows*, Amsterdam: North-Holland Publishing Company.

[173] Lopez, R.S. (1956), "The Evolution of Land Transport in the Middle Ages", *Past and Present*, 9(1), pp. 17-29.

[174] López-Córdova, J.E. and C.M. Meissner (2003), "Exchange-rate Regimes and International Trade: Evidence from the Classical Gold Standard Era", *The American Economic Review*, 93(1), pp. 344-353.

[175] Lugovskyy, V. and A. Skiba (2015), "How Geography Affects Quality", *Journal of Development Economics*, 115(C), pp. 156-180.

[176] Margary, I.D. (1973), *Roman Roads in Britain*, John Baker Publishing, London.

[177] Markusen, J.R. and A.J. Venables (2007), "Interacting Factor Endowments and Trade Costs: A Multi-Country, Multi-Good Approach to Trade Theory", *Journal of International Economics*, 73(2), pp. 333-354.

[178] Martin, W. and C.S. Pham (2008), "Estimating the Gravity Equation when Zero Trade Flows are Frequent", MPRA Working Paper 9453, University Library of Munich.

[179] Martin, W. and C.S. Pham (2015), "Estimating the Gravity Model When Zero Trade Flows Are Frequent and Economically Determined", Policy Research Working Paper 7308.

[180] McCallum, T. (1995), "National Borders Matter: Canada-U.S. Regional Trade Partners", *The American Economic Review*, 85(3), pp. 615-623.

[181] McCormick, M. (2001), *Origins of the European Economy: Communications and Commerce, AD 300-900*, Cambridge: Cambridge University Press.

[182] McCormick, M., G. Huang, G. Zambotti and J. Lavash (2013), "Roman Road Network (version 2008)", DARMC Scholarly Data Series, Data Contribution Series 2013-5.

[183] McCulloch, N. (ed.) (2009), *Rural Investment Climate in Indonesia*, Institute of Southeast Asian Studies: Singapore.

[184] Melitz, M.J. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica*, 71(6), pp. 1695-1725.

[185] Melitz, M.J. and G. Ottaviano (2008), "Market Size, Trade, and Productivity", *Review of Economic Studies*, 75(1), pp. 295-316.

[186] Michaels, G. and F. Rauch (2014), “Resetting the Urban Network: 117-2012”, CEPR Discussion Paper 9760.

[187] Miguel, E. and G. Roland (2011), “The Long-run Impact of Bombing Vietnam”, *Journal of Development Economics*, 96(1), pp. 1-15.

[188] Mimura, N., L. Nurse, R. McLean, J. Agard, L. Briguglio, P. Lefale, R. Payet and G. Sem (2007), “Small Islands”, *Climate Change*, pp. 687-716.

[189] Miroudot, S., J. Sauvage and B. Shepherd (2013), “Measuring the Cost of International Trade in Services”, *World Trade Review*, 12(4), pp. 719-735.

[190] Moenius, J. (2004), “Information Versus Product Adaptation: The Role of Standards in Trade”, Working Paper, Kellogg School of Management, Northwestern University, mimeo.

[191] Moïsé, E. and F. Le Bris (2013), “Trade Costs - What Have We Learned?: A Synthesis Report”, *OECD Trade Policy Papers*, No. 150, OECD Publishing.

[192] Montesquieu, C. (1750), *The Spirit of the Laws*.

[193] Montiel Olea, J.L. and C.E. Pflueger (2013), “A Robust Test for Weak Instruments”, *Journal of Business and Economic Statistics*, 31(3), pp. 358-369.

[194] Montolio, D. and A. Solé-Ollé (2009), “Road Investment and Regional Productivity Growth: The Effects of Vehicle Intensity and Congestion”, *Papers in Regional Science*, 88(1), pp. 99-118.

[195] Morgan, M.G. (1980), “Catullus and the Annales Volusi”, Quaderni Urbinati di Cultura Classica, New Series, 4, pp. 59-67.

[196] Mundlak, Y. (1978), “On the Pooling of Time Series and Cross-section Data” *Econometrica*, 46(1), pp. 69-85.

[197] Musti, D. (1990), *La Spinta Verso il Sud: Espansione Romana e Rapporti "Internazionali"*, in Arnaldo Momigliano; Aldo Schiavone (edited by), Storia di Roma Vol. I, Torino, Einaudi.

[198] Myrdal, G. (1968), *Asian Drama: An Inquiry into the Poverty of Nations*, New York: Twentieth Century Fund.

[199] Nagar, A.L. (1959), “The Bias and Moment Matrix of the General K-Class Estimators of the Parameters in Simultaneous Equations”, *Econometrica*, 27, pp. 575-595.

[200] North, D.C. (1981), *Structure and Change in Economic History*, Norton, New York.

[201] North, D.C. (1987), "Institutions, Transaction Costs and Economic Growth", *Economic Inquiry*, 25(3), pp. 419-428.

[202] North, D.C. (1989), "Institutions and Economic Growth: An Historical Introduction", *World Development*, 17(9), pp. 1319-1332.

[203] Novy, D. (2013), "Gravity Redux: Measuring International Trade Costs with Panel Data", *Economic Inquiry*, 51(1), pp. 101-121.

[204] Nunn, N. (2009), "The Importance of History for Economic Development", *Annual Review of Econonomics*, 1(1), pp. 65-92.

[205] Nunn, N. (2012), "Culture and the Historical Process", *Economic History of Developing Regions*, 27(S1), pp. 108-126.

[206] Nunn, N. and D. Puga (2012), "Ruggedness: The Blessing of Bad Geography in Africa", *The Review of Economics and Statistics*, 94(1), pp. 20-36.

[207] O'Rourke, K.H. and J.G. Williamson (1999), *Globalization and History: The Evolution of a Nineteenth-Century Atlantic Economy*, The MIT Press.

[208] Obstfeld, M. and K. Rogoff (2000), "The Six Major Puzzles in International Macroeconomics. Is There a Common Cause?", in B.S. Bernanke and K. Rogoff (eds.), *NBER Macroeconomics Annual*, 15, pp. 339-390.

[209] OECD/WTO (2015), "Aid for Trade at a Glance 2015: Reducing Trade Costs for Inclusive, Sustainable Growth", *OECD Publishing*, Paris.

[210] Olsson, O. and D.A. Jr. Hibbs (2005), "Biogeography and Long-Run Economic Development", *European Economic Review*, 49(4), pp. 909-938.

[211] Pascali, L. (2014), "The Wind of Change: Maritime Technology, Trade and Economic Development", University of Warwick, Department of Economics working papers.

[212] Pascali, L. (2016), "Banks and Development: Jewish Communities in the Italian Renaissance and Current Economic Performance", *The Review of Economics and Statistics*, 98(1), pp. 140-158.

[213] Pawson, E. (1977), *Transport and Economy: The Turnpike Roads of Eighteenth Century Britain*, London, Academic Press.

[214] Percoco, M. (2009), "History, Geography and Economies of Agglomeration: Evidence from Italian Cities", mimeograph, Università commerciale Luigi Bocconi.

[215] Perri, G. (2016), *Brindisi nel Contesto della Storia*, Lulu.com.

[216] Pflueger, C.E and S. Wang (2015), “A Robust Test for Weak Instruments in Stata”, *Stata Journal*, 15(1), pp. 216-225.

[217] Pinna, A.M. and V.M. Licio (2013), “Measuring Insularity as a State of Nature. Is There a Case of Bad Geography?”, CRENoS WP 2013/22.

[218] Pirenne, H. (1969), *Medieval Cities, Their Origins and the Revival of Trade*, Princeton NJ: Princeton University Press.

[219] Pomfret, R. and P. Sourdin (2010), “Why Do Trade Costs Vary?”, *Review of World Economics*, 146(4), pp. 709-730.

[220] Poulter, J. (2010), *The Planning of Roman Roads and Walls in Northern Britain*, Amberley Publishing.

[221] Puhani, P. (2000), “The Heckman Correction for Sample Selection and Its Critique”, *Journal of Economic Surveys*, 14(1), pp. 53-68.

[222] Putnam, R.D, R. Leonardi and R.Y. Nanetti (1993), *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton University Press.

[223] Quilici, L. (1989), *Via Appia. Da Porta Capena ai Colli Albani*, Fratelli Palombi Editori.

[224] Quilici, L. and S. Quilici Gigli (2004), *Carta Archeologica e Ricerche in Campania, Parte 3*, L'Erma di Bretschneider.

[225] Raballand, G. (2003), “Determinants of the Negative Impact of Being Landlocked on Trade: An Empirical Investigation Through the Central Asian Case”, *Comparative Economic Studies*, 45(4), pp. 520-536.

[226] Raftery, A.E., D. Madigan and J.A. Hoeting (1997), “Bayesian Model Averaging for Linear Regression Models”, *Journal of the American Statistical Association*, 92(437), pp. 179-191.

[227] Ramcharan, R. (2009), “Why an Economic Core: Domestic Transport Costs?”, *Journal of Economic Geography*, 9(4), pp. 559-581.

[228] Rappaport, J. and J.D. Sachs (2003), “The United States as a Coastal Nation”, *Journal of Economic Growth*, 8(1), pp. 5-46.

[229] Redding, S.J., D.M. Sturm and N. Wolf (2011), “History and Industry Location: Evidence from German Airports”, *The Review of Economics and Statistics*, MIT Press, 93(3), pp. 814-831.

[230] Richard, C.J. (2010), *Why We're All Romans: The Roman Contribution to the Western World*, Rowman & Littlefield Publishers.

[231] Riley, S.J., S.D. DeGloria and R. Elliot (1999), “A Terrain Ruggedness Index that Quantifies Topographic Heterogeneity”, *Intermountain Journal of Sciences*, 5(1-4), pp. 23–27.

[232] Roberts, H.V. (1965), “Probabilistic Prediction”, *Journal of the American Statistical Association*, 60(309), pp. 50-62.

[233] Rose, A.K. (2000), “One Money, one Market: The Effect of Common Currencies on Trade”, *Economic Policy*, 15(30), pp. 8-45.

[234] Rosenstein, N. and R. Morstein-Marx (2006), *A Companion to the Roman Republic*, Wiley-Blackwell.

[235] Roth, J.P. (1999), *The Logistics of the Roman Army at War (264 B.C.-A.D. 235)*, Columbia Studies in the Classical Tradition, 23, Brill.

[236] Rubin, J. and B. Tal (2008), “Will Soaring Transport Costs Reverse Globalization?”, *CIBC World Markets StrategEcon*, 27, pp. 4-7.

[237] Sachs, J.D. (2001), “Tropical Underdevelopment”, NBER Working Paper Series 8119, National Bureau of Economic Research Inc.

[238] Sachs, J.D. (2003), “Institutions Don't Rule: Direct Effects of Geography on Per Capita Income”, NBER Working Paper 9490, National Bureau of Economic Research Inc.

[239] Saing, C.H. and H. Kazianga (2017), “The Long-term Impact of U.S. Bombing on Education, Earning, Health, Fertility and Marriage in Cambodia”, mimeo.

[240] Sala-i-Martin, X. (1997), “I Just Ran 2 Million Regressions”, *The American Economic Review*, 87(2), pp. 178–183.

[241] Samuelson, P.A. (1954), “The Transfer Problem and the Transport Costs II: Analysis of Effects of Trade Impediments”, *The Economic Journal*, 64(254), pp. 264-289.

[242] Santos Silva, J. and S. Tenreyro (2006), “The Log of Gravity”, *The Review of Economics and Statistics*, 88(4), pp. 641-658.

[243] Schott, P.K. (2004), “Across-Product versus Within-Product Specialization in International Trade”, *The Quarterly Journal of Economics*, 119(2), pp. 647-678.

[244] Shepherd, B. (2010), “Trade Costs and Facilitation in APEC and ASEAN: Delivering the Goods?”, *Studies in Trade and Investment*, in Rising Non-Tariff Protectionism and Crisis Recovery, United Nations Economic and Social Commission for Asia and the Pacific (ESCAP), chapter 5, pp. 93-110.

[245] Sigelman, L. and L. Zen (1999), "Analyzing Censored and Sample-Selected Data with Tobit and Heckit Models", *Political Analysis*, 8(2), pp. 167-182.

[246] Smith, A. (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations*.

[247] Smith, W. (1890), *A Dictionary of Greek and Roman Antiquities*, London: John Murray, 1890.

[248] Snedden, R. (1998), *Technology in the Time of Ancient Rome*, Austin, Tex.: Raintree Steck-Vaughn.

[249] Soloaga, I. and L.A. Winters (2001), "Regionalism in the Nineties: What Effect on Trade?", *North American Journal of Economics and Finance*, 12(1), pp. 1-29.

[250] Sommer, M. (2015), "OIKOYMENH: Longue Durée Perspectives on Ancient Mediterranean 'Globality'" in Globalisation and the Roman World: World History, Connectivity and Material Culture, ed. Martin Pitts and Miguel John Versluys, New York: Cambridge University Press, pp. 175-197.

[251] Spiegelhalter, D.J., A.P. Dawid, S.L. Lauritzen and R.G. Cowell (1993), "Bayesian Analysis in Expert Systems", *Statistical Science*, 8(3), pp. 219-247.

[252] Spolaore, E. and R. Wacziarg (2013), "How Deep Are the Roots of Economic Development?", *Journal of Economic Literature*, 51(2), pp. 325-369.

[253] Staiger, D. and J.H. Stock (1997), "Instrumental Variables Regression with Weak Instruments", *Econometrica*, 65(3), pp. 557-586.

[254] StataCorp (2013), *Stata: Release 13. Statistical Software*, College Station, TX: StataCorp LP.

[255] Stock, J.H. and M. Yogo (2005), "Testing for Weak Instruments in Linear IV Regression", in Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg, New York: Cambridge University Press, pp. 80-108.

[256] Swann, P. (2010), "International Standards and Trade: A Review of the Empirical Literature", OECD Trade Policy Working Paper 97, OECD Publishing.

[257] Tabellini, G. (2010), "Culture and Institutions: Economic Development in the Regions of Europe", *Journal of the European Economic Association*, 8(4), pp. 677-716.

[258] Thompson, L. (1997), "Roman Roads", *History Today*, 47(2), pp. 21-28.

[259] Tobin, J. (1958), "Estimation of Relationships for Limited Dependent Variables", *Econometrica*, 26(1), pp. 24-36.

[260] Tsunokawa, K. (1983), "Evaluation and Improvement of Road Construction Cost Models", Unpublished Draft Working Paper, Washington, DC: World Bank.

[261] Turkson, F.E. (2012), "Using Observable Trade Data to Measure Bilateral Trade Costs in Sub-Saharan Africa", CREDIT Research Paper 12/06.

[262] UNCTAD (2009), *Review of Maritime Transport 2009*, United Nations, Geneva.

[263] UNCTAD (2016), *Review of Maritime Transport 2016 - The Long-Term Growth Prospects for Seaborne Trade and Maritime Businesses*, United Nations, Geneva.

[264] Valencia Caicedo, F. (2014), "The Mission: Economic Persistence, Human Capital Transmission and Culture in South America", mimeo.

[265] Volpe Martincus C., J. Carballo and A. Cusolito (2014), "Routes, Exports, and Employment in Developing Countries: Following the Trace of the Inca Roads", mimeo.

[266] Von Hagen, V.W. (1967), *The Roads That Led to Rome*, The World Publishing Company, Cleveland and New York.

[267] Wahl, F. (2015), "The Long Shadow of History: Roman Legacy and Economic Development - Evidence from the German limes", Hohenheim Discussion Papers in Business, Economics and Social Sciences, 08-2015.

[268] Waldinger, F. (2016), "Bombs, Brains, and Science: The Role of Human and Physical Capital for the Creation of Scientific Knowledge", *The Review of Economics and Statistics*, 98(5), pp. 811-831.

[269] Wei, S. (1996), "Intra-National versus International Trade: How Stubborn are Nations in Global Integration?", NBER Working Paper No. 5531.

[270] Welfare, H. and V. Swan (1995), *Roman Camps in England: The Field Archaeology*, London.

[271] Witcher, R. (2015), "Globalisation and Roman Cultural Heritage" in Globalisation and the Roman World: World History, Connectivity and Material Culture, ed. Martin Pitts and Miguel John Versluys, New York: Cambridge University Press, pp. 198-222.

[272] Wooldridge, J.M. (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press.

[273] Wooldridge, J.M. (2006), *Introductory Econometrics: A Modern Approach*, Third Edition, Thomson South-Western.

[274] World Bank (2002), *Globalization, Growth and Poverty: Building an Inclusive World Economy*, Oxford University Press.

[275] World Bank (2016), *Doing Business 2016: Measuring Regulatory Quality and Efficiency*, Washington, DC: World Bank.

[276] World Economic Forum (2015), *The Global Competitiveness Report 2015-2016*, Geneva: World Economic Forum.

[277] Xiong, B. and J.C. Beghin (2011), "Disentangling the Demand-enhancing Effect and Trade-cost Effect of Technical Measures in Agricultural Trade among OECD Countries", Proceedings Issues, 2011: Agricultural Price Volatility, Trade Policy and Food Security in Developing Countries, December 2011, St. Petersburg, FL, No. 116898, from International Agricultural Trade Research Consortium.

[278] Yaylaci, O. (2013), "Evolution of Trade Costs", Suffolk University mimeo.

[279] Zellner, A. (1986), "On Assessing Prior Distributions and Bayesian Regression Analysis with g-Prior Distributions", in Bayesian Inference and Decision Techniques: Essays in Honour of Bruno de Finetti, eds. P.K. Goel and A. Zellner, Amsterdam: North-Holland, pp. 233-243.