POLITECNICO DI MILANO

Facoltà di Ingegneria dei Sistemi



POLO REGIONALE DI COMO

Master of Science in
Management, Economics and Industrial Engineering

International Trade Costs and Transportation Modes:
The Case of Europe

Supervisor: Prof. Lucia TAJOLI

Master Graduation Thesis by:

Francesco BIANCHI, 751186

Sara POZZI, 748466

Academic Year 2010/2011

Abstract

This thesis explores the importance of product and country factors over the choice of transportation mode. In particular we focus on the synergic effect between the relevant determinants: products' price, delivery distance, time sensitiveness and transportation fares. Our goal is to determine the role of the main influencing factors in transportation pattern in order to identify their effects on international trade. This work addresses some of the most debated questions about the current international trade: *"Has distance died?";"What is the role of time in the current globalized context?";"What is the effective weight of transportation costs?";"Are the infrastructures relevant for international trade?"*.

Our results give possible explanations based on the European context. If markets present time constraints, air shipping usage should be higher; this is true only for very valuable goods while air shipping usage is actually lower considering cheap products: this is an evidence that transportation patterns and trade patterns are closely related, since low-value products have higher level of substitutability, and time constraints push towards high bilateral trade with close countries.

In general, any transportation choice is determined by the trade-off between certain "*extra costs*" and probable "*premium price*" for fast delivery: airplanes are the fastest but also the most costly transportation mean. Transportation "cost component" depending on distance doesn't seem very relevant, while delivery frequency and volume of trade can imply strong costs barrier. Indeed, the range of products traded by air is usually wider for longer distances, because time advantage seems to grow faster than "*extra costs"* with distance.

Distance is “dying” from a monetary costs’ perspective but it’s still puzzling international economics. On the one hand, transportation means are faster and faster and “fast means” are more and more affordable; on the other hand, the concept of speed and the benefit of time are something not really definable in absolute terms, since in the nowadays customer-oriented scenario, trade is timely only if it’s timely to the customer.

Keywords: Transportation costs, Distance, Globalization, Time, Air Shipping

Abstract

Questa tesi indaga sull'impatto di fattori di prodotto e di paese sulla scelta del mezzo di trasporto. In particolare, ci focalizziamo sulle sinergie che si instaurano tra le determinanti più importanti quali il prezzo, la distanza, la sensibilità della merce al tempo e il costo del trasporto. L'obiettivo è di definire il ruolo di questi fattori al fine di identificare il loro effetto sul commercio internazionale. Questo lavoro cerca di rispondere ad alcuni dei quesiti più dibattuti riguardo l'attuale scenario internazionale: "*La distanza è morta?";"Quale è il ruolo del tempo?";"Quanto sono rilevanti i costi del trasporto?";"Quale l'impatto delle infrastrutture?"*.

Se un mercato presenta vincoli di tempo, il trasporto aereo dovrebbe essere maggiore ma in realtà questo è vero solo per prodotti di valore. Questa è la prova di come le dinamiche di trasporto influenzino la "geografia" del commercio internazionale e viceversa: prodotti economici presentano di solito un maggior livello di sostituibilità e un eventuale limite di tempo si traduce in intensi scambi commerciali tra paesi confinanti.

Ogni decisione sulla scelta del mezzo di trasporto si basa su un forte trade-off: il trasporto aereo implica da una parte "*extra costs*" sicuri, dall'altra "*premium price*" probabili. Le voci di costo che variano in funzione della distanza non sembrano avere un ruolo rilevante, mentre la frequenza di consegna e il volume di scambio possono comportare "barriere di costo" notevoli. Infatti, all'aumentare della distanza, vantaggi di tempo crescono più velocemente degli "*extra costs*" richiesti, cosicché la gamma di prodotti commerciabile via aereo è più ampia se la distanza di consegna è maggiore.

La distanza "sta morendo" da un punto di vista dei costi, ma assume un ruolo importante e di difficile interpretazione per gli equilibri economici mondiali. Da una parte, tutti i mezzi di trasporto sono sempre più veloci e gli aerei sono sempre più accessibili; dall'altra, nello scenario attuale così fortemente orientato

al cliente, il concetto di velocità e di tempo è qualcosa di difficilmente quantificabile e lo scambio commerciale è tempestivo solo se è percepito come tale dal cliente.

Keywords: Costi del Trasporto, Distanza, Globalizzazione, Tempo, Trasporto Aereo

Index

1. INTRODUCTION 1

2. TRADE CONTEXT DESCRIPTION 4

2.1 The Era of International Trade 4

2.2 World Trade Growth: Influencing Factors 6

2.3 Trade Costs 10

2.4 The Role of Time 17

3. GOODS TRANSPORTATION 22

3.1 Transport Mode Decisional Process 22

3.2 How Goods Move 24

3.2.1 Sea Shipping 25

3.2.2 Air Shipping 27

3.2.3 Road Shipping 29

3.2.4 Rail Shipping 30

3.2.5 Transport Modes Classification 30

3.3 Transport Innovation and New Technologies 31

3.3.1 Sea Transport Innovation 32

3.3.2 Air Transport Innovation 35

3.4 International Agreements and Air Transport Costs 36

3.5 Transport Fares: Empirical Evidence 39

3.5.1 Distance 40

3.5.2 Weight and Volume 46

3.6 Focus on European Trade 52

4. TRANSPORTATION PATTERNS DETERMINANTS 64

4.1 Model Purpose 64

4.2 Artsana S.p.A. Interview 66

4.3 Trade Costs "Drivers" 73

4.3.1 Product 73

4.3.2 Time 77

4.3.3 Distance 83

4.3.4 Country Connectivity 89

4.4 Data Gathering, Statistical Adjustments and Dataset Creation 93

4.4.1 Point of Reference 94

4.4.2 Dependent Variable 95

4.4.3 Independent Variables 96

4.4.4 Outliers management 111

5. EMPIRICAL ANALYSIS: REGRESSION MODELING 113

5.1 Correlation Analysis 113

5.2 Regression Analysis 119

5.2.1 STEP (1): Multiple Linear Regressions 120

5.2.2 STEP (2): Non Linear Regressions 126

5.2.3 STEP (3): Regressions with Variables Interdependencies 133

5.2.4 Scenarios Analysis 160

5.2.5 STEP (4): Model Improvements 166

5.2.6 Multiple Interdependencies Analysis 179

5.3 Robustness Analysis 207

6. CONCLUSIONS 210

BIBLIOGRAPHY 220

APPENDIXES 229

APPENDIX A 229

APPENDIX B **232**

APPENDIX C **238**

APPENDIX D **246**

APPENDIX E **254**

APPENDIX F **257**

APPENDIX G **264**

List of Tables

Table 1 - World Trade as % of GDP 7
Table 2 – Standard Containers 26
Table 3 - Ship Types 27
Table 4 - Airbus 330 and Boeign 747 characteristics 28
Table 5 - Modes of Transport Ranking 30
Table 6 - Nautical Distance and Air Distance 40
Table 7 – Fares and Distance analysis 43
Table 8 - Variable Costs by Mode of Transport 44
Table 9 - Fares Difference 45
Table 10 - Air Fares depending on Distance 48
Table 11 - Fares depending on Volume 49
Table 12 – Fares Difference according to Volume 50
Table 13 - Countries Ranking by Export/Import 53
Table 14 - Main Product Categories Imported by EU27 59
Table 15 - Main Product Categories Exported by EU27 60
Table 16 - European Imports from Australia 62
Table 17 - European Import from Russia 63
Table 18 - Fixed and Variable Costs by Mode of Transport 85
Table 19 - Cost Structure of "Sea Shipping" 85
Table 20 - Time Difference 87
Table 21 - Coventional and Containerized Ports 91
Table 22 - Qualitative Scale for Time Sensitiveness 101
Table 23 – “Fare Difference (1)” Definition 106
Table 24 - "Fare Difference (2)" Definition 107
Table 25 - Weight of 50 dm^3 depending on Density 108
Table 26 - Correlation between Y% and all the Independent Variables 114
Table 27 - Correlation among Independent Variables 118
Table 28 - Linear Regressions (a) 120

Table 29 - Linear Regressions (b) 121
Table 30 - Linea Regressions (c) 122
Table 31 - Linear Regressions (d) 123
Table 32 - Linear Regressions with Youtliers% and Yspeed% 125
Table 33 - "Distance" Non-Linear Functions 127
Table 34 - "Density" Non-Linear Functions 127
Table 35 - "Airports Density" Non-Linear Functions 128
Table 36 - "Paved Roads" Non-Linear Function 128
Table 37 - "GNI/capita" Non-Linear Functions 129
Table 38 - "Time Difference" Non-Linear Functions 129
Table 39 - "Fares Difference (1)" Non-Linear Functions 130
Table 40 - "Fares Difference (2)" Non Linear Functions 130
Table 41 - "Value/Weight" Non Linear Functions 131
Table 42 - Multiple Non-Linear Regressions 132
Table 43 - Multiple Regressions with “Bulky” and “Density” Interdependences 135
Table 44 - Simple Linear Regression with "Fares(1)" and "Time Difference" 137
Table 45 - Multiple Regression with "Time Difference/Distance" and "Timeliness" Interdependency (a) 139
Table 46 - Multiple Regression with "Time Difference/Distance" and "Timeliness" Interdependency (b) 142
Table 47 - Simple Regression with "Distance-Timeliness" Interdependency 144
Table 48 - Simple Regresion with "Distance 0- Time Sensitiveness" Interdependency 145
Table 49 – Simple Regressions with “Fares/Time Difference” and “Distance” 146
Table 50 - Simple Regressions with "Distance/Time Difference" and "Density" 147
Table 51 - Simple Regressions with "Distance" and "Value/Weight" Interdependency 148
Table 52 - Simple Regressions with "Distance" (linear) and "Value/Weight"(linear) Interdependency 149
Table 53 - Simple Regressions with "Distance" (linear) and "Value/Weight"(logarithmic) Interdependency 149
Table 54 – Simple Regression with “Distance 0” and “Value/Wieght” Interdependency 152

Table 55 - Multiple Regression with “Value/weight –Distance” and “Timeliness-Distance” Interdependency 153
Table 56 – Simple Regression with “Time Difference” and “Value/Weight” Interdependency 154
Table 57 - Simple Regression with "Timeliness" and "Value/Weight" Interdependency 155
Table 58 - Simple Regression with "Value/Weight" and "Density" Interdependency 157
Table 59 - Comple Multiple Regressions with Interdependencies (a) 159
Table 60 – “Fares Difference (3)” Non-Linear Functions 167
Table 61 - Simple Regressions including Inderdependencies with "Fare Difference (3)" 168
Table 62 - "Air Shipping" probability depending on Value/Weight 170
Table 63 – Quartiles 170
Table 64 - Multiple Regressions with all of the"Fares Difference 3" Interdependencies 172
Table 65 - Complete Multiple Regressions - With Country Dummy (a) 177
Table 66 - Complete Multiple Regressions - With Country Dummy (b) 178
Table 67 - Y% according to V/W and "Fare Difference (3)" 180
Table 68 - Quartile Values 183
Table 69 - Scenarios Analysis According to Different Variables Quartiles 183
Table 70 - V/W Thresholds (1) 185
Table 71 - Y% Slacks 188
Table 72 - Y% Slacks 188
Table 73 - Quartiles Values 191
Table 74 - Y% = f(V/W; Fare3; Time Sensitiveness; Distance) 192
Table 75 - Value/Weight Tresholds 193
Table 76 - V/W Thresholds Values (Y=0.3) 201
Table 77- Thresholds V/W --> Y=0 205
Table 78- Thresholds V/W --> Y=30% 205
Table 79 - Probit Regressions 209

List of Figures

Figure 1 - International Trade Growth 4
Figure 2 - Evolution of Companies Objectives....19
Figure 3 - Costs and Service Level 23
Figure 4 - European Trade by Mode of Transport....32
Figure 5 - Worldwide Fleer by Type of Ship....34
Figure 6 - "Sea" and "Air" Routes from Livorno to New York 41
Figure 7 - "Sea" and "Air" Routes from Rotterdam to Melbourne....41
Figure 8 - "Sea" and "Air" Routes from Genoa to Tokyo....42
Figure 9 - Air Shipping Fare depending on Distance 43
Figure 10 - Sea Shipping Fare depending on Distance....44
Figure 11 - Fares Difference depending on Distance 46
Figure 12 - Air Fares depending on Weight....48
Figure 13 - Fare and Volume Relationship....50
Figure 14 - Fares Difference according to Volume....51
Figure 15 - Countries Trade by Mode of Transport 55
Figure 16 - Total European Import (in value) from Different Countries 56
Figure 17 - Total European Import (in value) from Different Countries by Mode of Transport....56
Figure 18 - Comparison between Total Value and Total Weight of European Import....58
Figure 19 - Value/Weight Comparison....58
Figure 20 - Value/Weight of European Import by Mode of Transport....74
Figure 21 - Worldwide FDI Stock....82
Figure 22 - "Step-wise"Transport Fare....83
Figure 23 - Transportation Cost Structure....84
Figure 24 - Cost Functions by Mode of Transport....86
Figure 25 - "Sea" and "Air" Shipping Time 88
Figure 26 - Italian and Turkish Investments in Port Infrastructures 90
Figure 27 - Freight Traffic at the World's Largest Airports, 2008 92
Figure 28 - Correlation between "Air Shipping" probability and Airports Density 115

Figure 29 - Correlation between "Air Shipping" probability and Value/Weight.....116
Figure 30 - Correaltion between "Yspeed%" and Value/Weight 116
Figure 31- Correlation between "Air Shipping" probability and Value/Weight 117
Figure 32 - Derivates' Functions Representation 152
Figure 33 - Y% =f (ln(V/W|Fare3;Timeliness) 184
Figure 34 - Y% =f (ln(V/W)|Range of Fare3; Time Sensitiveness) 186
Figure 35 - Y% =f (ln(V/W)|Interval of Fare3; Time Sensitiveness).....186
Figure 36 - Y% "Non-Timely Scenario" 190
Figure 37 - Y% "Timely Scenario" 191
Figure 38 - Thershold Vaues/Weights = f(Distance) 194
Figure 39 - ln(Thresholds V/W)=f(Distance) --> for Y=0 and Y=0.3 (Low Fare3) 196
Figure 40 - ln(Thresholds V/W)=f(Distance) --> for Y=0 and Y=0.3 (Medium Fare3) 197
Figure 41 - ln(Thresholds V/W)=f (Distance) --> for Y=0 and Y=0.3 (High Fare3) 197
Figure 42 - V/W = f (Distance) --> "Non-Timely Scenario" 199
Figure 43 - V/W = f (Distance) --> "Timely Scenario" 199
Figure 44 - V/W = f (Distance) --> "Non-Timely Scenario" 200
Figure 45 - V/W = f (Distance) --> " Timely Scenario" 200
Figure 46 – Y%=f (ln V/W|Time Difference; Fare3) → "Non-Timely Scenario" 202
Figure 47 – Y%=f (ln V/W|Time Difference; Fare3) → "Timely Scenario" 203
Figure 48 - Thresholds V/W =f (Time Difference) → Case:Y=0% 204
Figure 49 – Transportation Pattern Determinants 212
Figure 50 - Transportation Pattern Interdependent Determinants 212
Figure 51 - Data Base (1) 246
Figure 52 - Data Base (2) 247

1. Introduction

Globalization has radically changed international trade, giving to "*trade costs*" a more and more important role for the economic balance of companies and countries. In literature, the theme of the globalization is widely discussed and many economists and journalists have stressed the link between global markets and the "dramatic" fall of transportation costs in the second half of the 20th century. However, the interpretation of the role played by transportation costs is still unsolved, so that some authors refer to the theme with the eloquent appellation of "distance puzzle". The point is that empirical studies have shown that distance coefficient in gravity model's applications is even more important today than in the past (despite the lower and lower transportation costs). So, the popular question remains open: "*What's the reason for the puzzling persistence of distance in international trade?*". Economic literature is plenty of different discussions and theories about it.

One of the most debated interpretations regards the role of time in trade. Distance is not only a "proxy" for costs but it also implies a certain time to deliver. Even if costs are not an insurmountable constraint anymore, distance can still be a barrier since it can cause delays in transactions completion.

Air shipping can represent a natural solution for distant trade to be "timely" and in the last decades, airplanes have been more and more used for cargo transportation. On the one hand, it can be the result of technological breakthrough and consequent lower "planes rates" (introductions of the "jet engine" and the so called "wide body" aircraft are only some of the

improvements made); on the other hand, it can be also the obvious consequence of the increasing need for fast delivery.

Although airplanes costs are going down, they still remain sensibly higher than sea or truck shipping costs, so that, no one should even think about delivering by air from a mere financial perspective. Nevertheless, the idiom "time is money" has never been as true as today and customers are more and more willing to pay "*premium price*" to get quickly what they need.

It's evident that we are dealing with a trade-off between expected premium prices and certain extra-costs.

In our thesis, we try to identify the main determinants for air shipping to be affordable and to define suitable indicators to measure these determinants. The idea is to collect data and evidences to assess the role played by the different variables over the air shipping probability, in order to better describe the features of the "*premium price*"-"*extra costs*" trade-off.

The final objective of our study is to give a contribution to the discussion about the role of distance, time and transportation costs in the current international trade, focusing on European commerce.

The thesis is structured as follows.

In Chapter 2 we go through a review of the economic literature concerning the most debated topics about international trade costs.

In Chapter 3 we describe the international transportation structure from a technical point of view on the basis of data coming from literature and logistic companies' websites (e.g. TNT, FedEx and DHL). In particular, we focus on technical features, technological developments, regulations and cost structures of the different transportation modes. Finally, we "take a picture" of the European trade context.

Starting with the case of Artsana S.p.A, we identify in Chapter 4 all of the relevant determinants to be tested. Then we present the indicators we use in regression; finally, we describe the database structure and the outliers management activities made to guarantee statistical consistency.

In Chapter 5 we show all the steps carried out to improve our regression model. We start testing linear relationships among variables, then we look for best non-linear functions and we analyze the interdependencies among the determinants. Finally, we check for the robustness of our model.

In Chapter 6 we conclude discussing the results.

2. Trade Context Description

2.1 The Era of International Trade

World trade increased a lot in the last decades (Figure 1). From 1990, international trade started to grow on average 9% annually. Evolution of developing economies, introduction of new technologies, establishment of multinational firms and the spread of supply chains all over the world brought the "Globalization[1]". In this context there is a need for continuous exchange of goods from place to place.

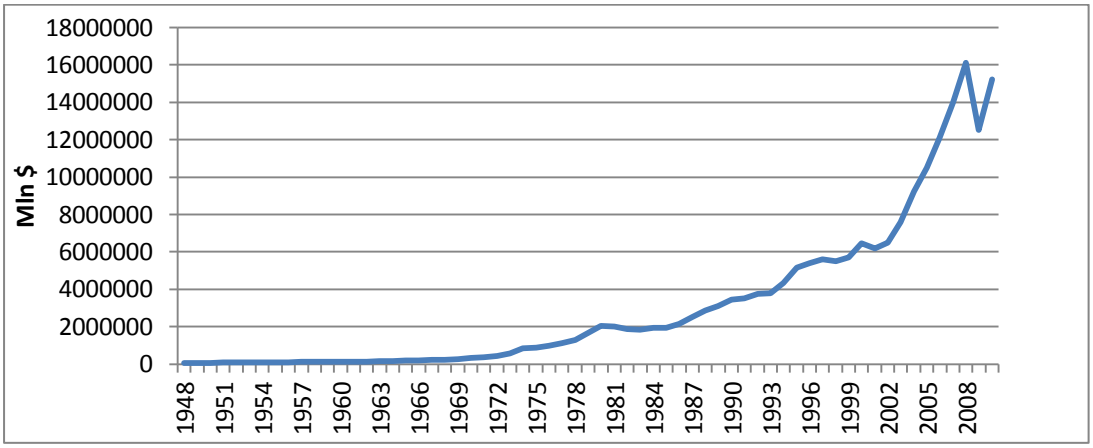


Figure 1 - International Trade Growth
Source: WTO

[1] In economic literature there are several definitions about "Globalization". A popular statement that we can embrace is provided by the economist Bhagwati (2004): "Economic globalization constitutes integration of national economies into the international economy through trade, direct foreign investment (by corporations and multinationals), short-term capital flows, international flows of workers, and flows of technology."

Karel De Gucht, European Commissioner for Trade, stated that "*Trade is an engine for global growth*" (2010, "Trade, Growth and World Affairs", European Commission Communication). Trade is "the heart" of international economy and studies and analyses about it, under different points of view, constitute a valuable knowledge about the future development of the world economy.

2.2 World Trade Growth: Influencing Factors

World trade has grown. But, which are the causes? Which are the factors that made possible international trade's expansion?

This issue has been debated several times by economists and journalists and they have proposed many different theses.

In 1995, Paul Krugman opened the following question:*"Why has world trade grown, and what are the consequences of that growth?"* (Krugman, 1995). He noted that journalists and economists tend to propose two different answers:

- Journalists use to indicate **technology innovation** as main driver: **lower transport costs**, more **speed**, more **reliability** and faster communication have boosted international trade.
- Economists tend to underline the relevance of **political factors**: the liberalization of international market carried the worldwide countries to trade with each other.

Krugman tries to answer to the question by proposing three main influencing factors.

1) **Trade policies** and **political actions** like **GATT** signature (1947), the introduction of the **WTO** (1995) and the definition of unilateral agreements by developing countries (e.g. **APEC** – Asia-Pacific Economic Cooperation, 1989) have promoted cooperation and trade among countries.

   Krugman proposes simple empirical evidences. Looking at the evolution of the world merchandise export in the long term (Table 1), we discover that trade had two "economic waves"; data show an increase in trade up

to 1913; a substantial drop-off between 1913 and the early post World War II years; and a recovery that has continued to the present.

YEAR	1850	1880	1913	1950	1973	1985	1993
World Trade as % of GDP	5,1	9,8	11,9	7,1	11,7	14,5	17,1

Table 1 - World Trade as % of GDP

Source: WTO

During war periods there was the implementation of strict protectionist barriers like tariffs, quotas and exchange controls. "*The correspondence between the time profile of trade policy and that of trade shares is prima facie evidence that political factors have played a major, perhaps dominant role in the growth of world trade since 1950"*(Krugman, 1995).

However, he observes also that market liberalization is not enough to produce full economic integration: other political issues could obstacle countries' cooperation. Evidence of this is the bilateral trade between the US and Canada. John McCallum has demonstrated that bilateral trade among far Canadian provinces is on average much higher than trade with closer American states: the "real slack" is twenty times higher than expected with gravity model equations. Canada and the US are very similar in terms of language, culture and institutions; both of them have good infrastructures and high GDP per capita; moreover, there are no significant trade barriers such as tariffs and quotas that could undermine the analysis. Evidently, the political boundaries themselves are an "effective obstacle" to trade.

2) Trade growth is the results of **changing in world output distribution**. The traditional gravity model is presented in the following form:

$$T_{ij} = \beta_0 + \beta_1 Y_i^{\alpha} + \beta_2 Y_j^{\beta} + \beta_3 D_{ij}^{-\gamma}$$

T_{ij} is the trade between country i and country j; Y is the country's GDP and D_{ij} is the distance between the two reference countries. GDP expresses simultaneously the capability of a country to produce (and export) goods and in the same time the country's funds availability to buy (and import) goods. The concept is: the more the countries' GDP are similar, the higher is the bilateral trade. Krugman (1995) demonstrates that "*the national distribution of gross product among the world's economies…has become somewhat more equal over time: the sum of squared shares fell from 0,13 in 1960 to 0,104 in 1993"* (see table 1). Countries are becoming more similar (the number of country-equivalents in the world economy has risen from 7,7 to 9,6). This income convergence could partially explain international trade growth.

3) **Technological change** is an evident evolution whose effects are difficult to be detected. Krugman states that "*world trade is not completely determined by technology: transportation and communication technologies were better in 1950 than they were in 1913, but the world economy was substantially less integrated*". He declares that technology is not a dominant factor in trade growth.

Many others economists have proposed explanations that are similar to Krugman's ones.

R.C.Feenstra (1998) states that trade liberalization, transport costs falling and convergence in economic size have boosted up international trade. Feenstra (1998) stresses also the role of **"disintegration" of production processes** all over the world: multinational firms and international outsourcing have increased the volume of international trade. Let's think about the popular "Barbie" doll produced by Mattel. Rone Tempest (famous American journalist) published in 1996 an article on the "Los Angeles Time" about the role of China in Barbie production; the article incipit was: "*A Barbie doll is for sale at the*

Anaheim Toys "R" Us store in a bright cardboard-and-cellophane box labeled "Made in China." The price is $9.99.”. The journalist stressed that the label “Made in” doesn’t represent the real origin of production and different production phases are scattered around the world. The raw materials for the doll (plastic and hair) come from Taiwan and Japan. Moulds and additional paints used in decorating the dolls come from the United States. China supplies only labour and cotton for the dresses.

Baier and Bergrand (2001) try to test empirically the relevance of the different influencing factors for trade growth. They verify “*the relative effects of transport costs reduction, tariff liberalization and income convergence* (and income growth) *on the growth of world trade among several OECD countries between the late 1950s and the late 1980s*”. They find that 67-69% of trade growth is explained by income growth, 23-26% by tariff-rate reduction and 8-9% by transport costs reduction; GPD convergence “virtually” doesn’t explain trade growth.

Other economists, instead, tend to give more relevance to **technologic improvements** and to the consequent **transportation costs falling**. David Hummels (2007) declared that “*One prominent possible explanation for the rise in international trade is a* ***decline in international transportation costs***”. Empirical evidences demonstrate a positive correlation between shipping costs reduction and trade growth in the first economic wave (Table 1, 1850-1913) between 1850 and 1913 (Estevaordal et al., May 2003). Anderson and Van Wincoop (2004) focus on the great relevance of trade costs, stating that “*Trade costs are large, even aside from trade-policy barriers and even between apparently highly integrated economies*”. They affirm also that direct policies instruments (like tariffs and quotas) are less relevant than other economic policies like infrastructures’ development, bureaucracy system and laws.

In conclusion, we can say that the fundamental question "*Why has world trade grown?*" remains intensively disputed; there are different points of view and it is difficult to deny a theory rather than another one.

Certain evidence is that aggregated **trade costs** (including tariff, transportation, bureaucracy, etc) **are decreasing over time** thanks to the interaction of technological improvements and trade liberalization.

2.3 Trade Costs

"*Globalization should generate a dispersion of economic activity reflecting a decline in transaction costs, especially transport costs. But studies based on the traditional gravity model of international trade—the workhorse for studies on the pattern of trade and the influence of transport costs—do not reach that conclusion*" (Brun et al., 2005).

In the globalized context, trade costs become more and more relevant since average distance of trade is bigger and bigger. On the one hand, direct shipping costs has lowered a lot in the last decades with new efficient innovations and technical solutions; on the other hand, average distance of trade raised with the globalization (farer and farer countries started being competitive in many industries).

According to Berthelon and Freund (2008), trade costs are on average nearly twice as large as production costs: they must be fully analyzed to understand trade patterns and comparative advantages in the nowadays scenario.

If transportation costs fall down, we would expect an increase in the overall international trade, and especially an increase in bilateral trade between far

partners (if we consider that transportation costs tend to increase with distance). According to Berthelon and Freund (2008), globalization brought a 300% increase of trade volumes but trade volumes kept being distance sensitive. In other words, trade grows proportionally to distance.

Most of the literature, according to the traditional gravity models, relies on distance as a proxy for transportation costs. Thus, reduction of transportation costs is supposed to reduce significantly trade elasticity to distance. The traditional managerial and economic models usually consider just "an aggregate of transport costs" to capture the penalty of distance, while simply remarking that these costs are shorthand for a wide range of penalties. The popular "iceberg model" was introduced by Samuelson in 1954. Literally "*A cost of transporting a good that uses up only some fraction of the good itself, rather than using any other resources. Based on the idea of floating an iceberg which is costless of the iceberg itself that melts. It is a very tractable way of modeling transport costs since it impacts no other market" (*Deardoffs' Glossary of International Economics[2]). The "iceberg" is the initial value of a good that melts during transportation losing a fraction of its value. Obviously, the longer the shipment the bigger the melting.

In the reality, gravity models analysis (done for several different years) shows that the absolute value of the "distance coefficient" in bilateral trade almost always increases over time, despite shipping costs reductions (Brun et al., 2005).

Even Leamer and Medberry (1993) analyze how distance coefficient didn't fall in the last twenty years, while Disdier and Head (2008) analyze the "impact of distance" on bilateral trade through a "meta-analyses investigation". Their paper goes through a comprehensive, quantitative analysis of the magnitude of the distance effect and the factors that explain variation in the estimates. The

[2] http://www-personal.**umich.edu**/~alandear/glossary/i.html

most striking finding is that distance effects decreased slightly between 1870 and 1950 and then began to rise. Evidently, the effect of distance on trade can't be summarized by transportation (freight) costs.

Finally, "*Grossman (1998) performs a simple calculation showing that estimated distance effects are about an order of magnitude too large to be explained by shipping costs*." (Disdier and Head, 2008): in other words, shipping is less costly today if compared with the past, but distance has a stronger and stronger impact on trade and consequently overall trade costs must be higher.

Duranton and Storper (2008) gave a first explanation to this phenomenon, explaining that trade costs are not proportional to transportation costs. For instance, lack of information about far partners and tariffs must be considered.

According to Harrigan and Venables (2004), trade costs include freight and other monetary transaction costs; lack of information about markets, suppliers and local institutions; difficulty in monitoring contracts; the impossibility of face to face contact and communication; and the fact that distance introduces delay into completion of trade. "*It is unlikely that summarizing these penalties as a single value of trade costs is adequate for understanding their effects*" (Harrigan, Venables, 2004).

So, lack of information and protectionism are the most intuitive thesis discussed in literature. However, the ICT breakthrough should have made relatively easier to trade with far partners and markets haven't been affected by sensible increases in regionalism and protectionist policies.

Noblet et al. (2010) list some of the most discussed theories about this topic. The first reason is about the interpretation of the ***"distance coefficient"*** in gravity models. Deardoff (1998) defined it as the product between elasticity of trade to trade costs and elasticity of trade costs to distance (the latter equal to ratio between marginal and average trade costs). If marginal costs fall proportionally with average costs, why should the ratio change? Coe et al. (2002) introduce the

so called “multilateral resistance term” in gravity model: if transport costs fall relatively more for closer countries, trade becomes more distance sensitive. However, Dias (2010) shows how transportation costs should have fallen more for long than for short distances undermining the previous thesis.

The second possible reason is about the changing ***world trade geography*** that could bias results. Globalization could have been caused by the increased world GDP rather than lower transportation costs (there’s no direct causal effect between trade and shipping cost) (Leamer and Levinsohn, 1995). Indeed, Berthelon and Freund (2008) show how frictionless DOT (theoretical distance of trade weighted on countries GDP) increased much more than actual DOT (distance of trade). This to explain that new developing countries are actually changing trade patterns and affecting distance coefficient (if a close developing country starts producing something, it “steals” trade shares to far partners. The coefficient distortion is the result of the “new trade geography”). Indeed, Disdier and Head (2008) show that the distance coefficient falls significantly over the last decades considering developed countries only.

However, the effect can’t be explained looking at trade geography only, since the registered 25% average reduction in shipping costs must have a strong and consistent impact (Hummels, 2007). As already introduced above, ***regionalism*** can have an impact on trade flows. For instance, the NAFTA institution increased trade among member countries (Coughlin, 2004).

As mentioned above, ***trade costs are not only shipping costs***. It’s necessary to look at them in a broader way, considering that different costs components can change over time. For Coe and others (2002) there can be omitted variables in the transport cost-functions commonly considered in economics.

Brun and others (2005), introduce a “remoteness coefficient” and an augmented transportation costs keeping into consideration indexes of infrastructure, price

of oil, and composition of trade. With these changes distance elasticity trend is negative (and the result is very strong and robust by clustering only reach countries, according to GDP pro-capita). Dias (2010) tries to classify costs component in "distance-related" (freight costs) and "non distance-related" (tariffs, FDIs as a proxy of decentralization barriers): if non distance-related costs lower, distance impact shouldn't lower.

Duranton and Storper (2008) and Noblet and others (2010) consider transportation costs not to be monotonic with "transfer cost" (their definition of trade cost). Indeed, shipping costs collapse implies birth or endowment of other less direct source of costs.

Duranton and Storper (2008) explain how these extra-costs are related to information flows. Lower shipping costs allow firm to trade higher quality products, this resulting in products that are more expensive to trade since transfer costs are considered an increasing function of quality (the idea behind it is that a larger portion of quality is lost during transportation, according to the "iceberg perspective"). Noblet and others (2010) deepen this perspective: lower shipping costs push firm towards greater complexity levels. The higher complexity is consequence of a larger variety of intermediate products to be purchased. In an "incomplete contract NNTT[3] context" (like the one assumed by Noblet and others), there is a gap between intermediate specifications and real products features. And this gap becomes higher for higher varieties. So, if transportation costs go down, paradoxically trade could be more sensitive to distance because of these coordination costs. A similar explanation is given by

[3] NNTT is the acronym of "New New Trade Theory". It's an evolution of the "New Trade Theory": the collection of economic models based on the assumptions of increasing return to scale and networks effects. NNTT introduces in the economic models the concept of having incomplete contracts because "downstream firm is likely to observe a distance between the optimal set of characteristics and the actual one" (Noblet et al., 2010).

Harrigan (2004). If it's quite easy to detect direct monetary costs for distance in trade (e.g. freight costs, tariffs, fuel), it's much more difficult to foresee and translate in numbers how much the uncertainty about markets, institutions and delivery times impact on profits and managerial decisions. According to Harrigan (2004): "*distant delivery results in higher lead-time uncertainty and consequently in production costs because of the lack of synchronization".*

The last reason is about the ***compositional effect***. Distance coefficient is an average result. If trade volumes for different industries have changed over time, the distance impact is not supposed to be consistent anymore. Moreover, the effect could be due to only a part of the industries that became more distance sensitive rather than to a structural increase of the trade elasticity to distance.
Coe and others (2002) give this kind of explanations: the "*exclusion of zero observations from the model could bias estimation*" because of the changing composition of trade. Berthelon and Freund (2008) define a model to assess compositional effect together with trade elasticity to distance at the industries level. Compositional effect doesn't seem to be strong, while distance coefficient is stable due to only a part of product categories (about the 40%) that became very distance sensitive over time. These categories are those with very high substitutability coefficient.

Moreover, the nowadays scenario is the one in which competition is stronger and stronger, and the customers, very demanding, have easy access to all of those information necessary to evaluate a companies' offer. High level of service and customization of products, quick time to market (product lifecycle is shorter and shorter) became the managerial challenge in the last decades and **timeliness** in trading for any stage of the macro-process has to be an important critical success factor (Harrigan, 2005).

Harrigan introduces the topic of the timeliness, a real source of competitive advantage today. Under this perspective, distance matters because the time for

deliveries matters. Airplanes can solve the distance puzzle, allowing far but timely shipping. In the appendix (Appendix A) we go through the timeliness model by Harrigan. He started from the Eaton-Kortum economic equilibrium.

Let's think about the high-tech industry; it doesn't make any sense to deliver computer chips in months shipping by sea to far destination. It would result in products discounting or technical obsolescence. Fast component delivery results in less costly production (companies can rely on actual customers order and not on imprecise forecast) and fast final products delivery results in higher service level and profits (Harrigan and Venables 2004).

So distance has a direct impact on transportation costs and an indirect impact on level of trade uncertainty and trade timeliness. The role of time is fundamental today, and it's going to be a stronger and stronger challenge in the future.

2.4 The Role of Time

"*Time is Money*"

This "truism" is the simple explanation given by Evans and Harrigan (2003) to justify the fact that some goods are shipped by air, although sea shipping is considerably cheaper.

A cargo ship leaving the Genoa harbour takes two weeks to get Tokyo. It takes only two days by plane. This time difference motivates the willingness-to-pay higher price for air shipment.

But, what is the real "cost of time"? Hummels (2001) tries to answer the question by considering **time as a trade barrier**. Time cost is determined by two main cost's voices:

- **Inventory-Holding costs** → They include both inventory cost and opportunity cost. Effective inventory costs are all those costs that are related to the physical holding of stock inventory (e.g. storage, insurance security, etc.). While opportunity cost represents the "missed returns" from the capital chained in stock inventory.

- **Depreciation** → Depreciation is the loss of value caused by the fact that a new product is preferable to an older one. There are many causes of depreciation; for example a daily newspaper losses completely its value the day after its publication; other products like foodstuff and cut flowers are perishable; High-Tech products risk to become obsolete in short time due to the continuous technological improvements and to the sudden fashion evolution.

Hummels (2001) tries to estimate the magnitude of time cost and he discovers that "*For manufactured goods I find each day in travel is worth an average of 0,8 percent of the value of the good per day, equivalent to a 16 % tariff for the*

average length of ocean shipment". He also indicated that "*each additional day in ocean transit reduces the probability that a country will export to the US by 1 percent (all goods) to 1,5 percent (manufactures goods)".*

Djankow et al. (2006), follow the Hummels theory trying to determine how much time delay affects international trade. They implemented a regression model discovering that "*one-day reduction in delays before a cargo sails to its export destination is equivalent to reducing the distance to trading partners by about 70 km*". Another formulation of their result is that "*on average each additional day of delay reduces trade by at least one percent*" (that is consistent with Hummels' result). They also find that delays have higher impacts on time-sensitive sectors such as perishable agricultural products.

The previous empirical evidences on time costs stress the growing importance of timeliness in international trade. But, which are the root-causes of this "time relevance"? Economic literature is "plentiful" of theories and explanations about this.

Demand volatility is one of the first causes. Customer's tastes and needs are changing and companies have to follow the customers' wants. In the past, firms offered a "mass market" production: they produced standardized goods in high volumes. In the first half of the twentieth century, the automaker giant Ford Motor Company, founded by Henry Ford in 1903, embodied the mass market paradigm; the Ford-T model (1908) was extremely standardized: it was the unique model for everybody; in 1909 Henry Ford remarked: *"Any customer can have a car painted any colour that he wants so long as it is black"* ("My life and work", chapter IV, pages 72-73, 1922). In such a context, it's clear that companies had no problems in demand forecasting and production planning: they knew what to produce and in what quantity.

Over the years customer needs have radically changed and companies had to modify their production paradigm. Philip Kotler (1989) states that "*It is my belief that the "mass market" is dead—segmentation has now progressed to the*

era of ***mass customization***". Hence, companies had to change their way to manage production in order to manufacture more product varieties and more customized goods in lower volumes. In addition to the increasingly complexity of product characteristics, customers changes their tastes faster than before; Deardorff (2001) states: "***Over the time with surprising speed, consumers change what they want****, and so, therefore, do the producers who seek to serve them. Goods that might have been just what they wanted one day become unwanted – or at least less than ideal – the next*". The result is that **product life cycle is shortening**, obsolescence risk is higher and holding inventory becomes more costly (and risky).

In the following image (Figure 2), there is the representation of companies' goals evolution: goals have changed according to the variation of customer requirements. Companies are now "customer oriented" (in the past the focus was on internal efficiency and cost) due to the higher dynamism and complexity of the environment. In this new context **time and coordination become critical success factors to exceed competitors**.

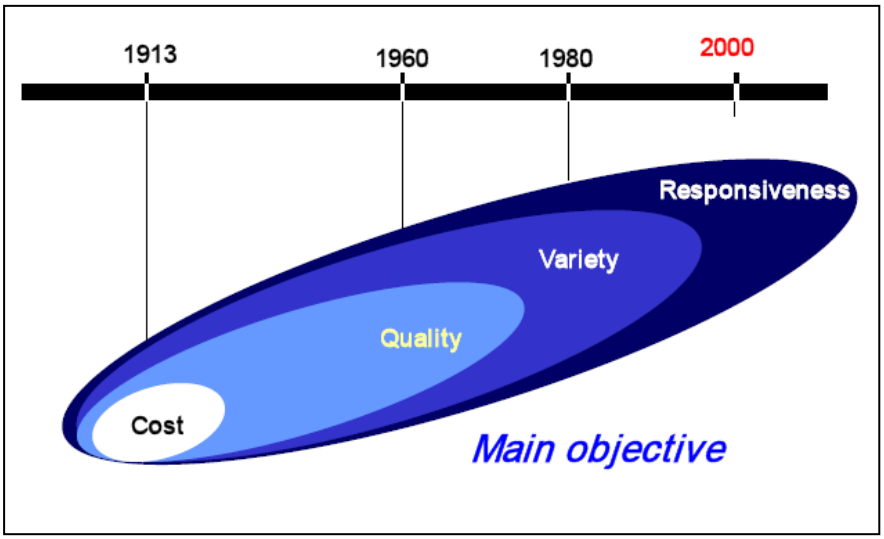


Figure 2 - Evolution of Companies Objectives

Source:Prof. Marco Taish from "Industrial Technology" course slides , 2010

In the last decade, this trend has been more and more stressed: customers don't adapt their needs to existing products anymore, while companies have to perfectly match customer changing needs with newer and newer products. Co-production and co-design are the "incarnation" of these current contingencies.

One of the adopted managerial solutions for facing this dynamic environment is the **Just-in-Time technique** (JIT); the basic JIT's assumption is to produce what the customer wants in the moment that he/she wants: in this way the inventories are eliminated and the related "obsolescence risk" is removed. Hutchins (1988) defines JIT in the following way: "JIT is *the term used to indicate that a process is capable of instant response to demand without the need for any overstocking, either in expectation of the demand being forthcoming or as a result of inefficiencies in the process*".
It's evident that **timeliness** has a crucial role in **JIT implementation**: all of the components must be available in the right quantity and in the right moment. A component delay stops the production process, causing idle time and reduction of customer satisfaction.

Time costs are also magnified by the current expansion of **vertical FDI and international outsourcing**. In case of vertical FDI, companies have a "direct" control on firm's production and facilities: companies spread different production phases around the world in order to exploit the advantage of the local resources. On the contrary, in case of international outsourcing, companies don't have direct control on the foreign firm (less risk but less control). **"Need for Timeliness"** is a common requirement for both of them; the production process is split around the globe and it can work if and only if **synchronization and timeliness** are fulfilled across the different productive phases. In case of international outsourcing, risk of delay could be higher due to the absence of control on firm's operations. Harrigan and Venables (2004) demonstrate that this "need for timelines" generates the tendency to cluster plants around the assembler or retailer.

Deardorff (2001) explores the importance of time in international trade, taking into accounts all the peculiarities of the current context. Deardorff, in compliance with Hummels (2001), consider the time cost as **depreciation**. He develops a model in which products depreciate over time, perhaps because of the characteristics of goods that consumers desire change randomly over time. "*If goods lose value quickly over time, then those who produce these goods must also move quickly, in both producing them and in delivering them to their users*". In his model, Deardorff demonstrates that time can be a relevant component for determining comparative advantage. Fast production and rapid transportation minimize products' depreciation; hence, countries that are able to produce and delivery in very short time could have a **comparative advantage** in time-sensitive goods (all those goods that depreciate faster than the others).

Y. Li and J.S.Wilson (2009) examine the role of time in international trade by testing "*whether firms of more time-sensitive sectors performs better in countries that are able to move goods quicker for export*". They implemented a "Difference-in-Difference" approach on a unique cross-country data set and they discover that "*firms in more time-sensitive industries are less likely to export in countries having longer required time for export to take place*". The two economists reach the same evidence proposed by Deardorff's model: the country's capability to move goods faster will determine its comparative advantage in time-sensitive industries.

3. Goods Transportation

3.1 Transport Mode Decisional Process

Transportation is a cost. According to a research by the "Commission of the European Communities", logistic costs represent 10-15% of unit costs for finished products and 60% of these logistics costs come from transportation activities. Tactical decisions about "modes of transportation" have a direct impact on companies' profitability and competitiveness (for instance by affecting final prices).

Transportation and logistics activities must be seen as strategic activities creating value to the customer: Ronald H. Ballou (R.H. Ballou, "'Logistics, Supply Chain and Transport Management", publication for Cambridge International College"), states that "*...logistics is about creating **value** - value for customers and suppliers of the firm, and value for the firm's stakeholders...";* actually, a product doesn't have any real value until it is delivered when and where the customer wants.

Transportation costs and service level are clearly in trade-off (Figure 3). Firms can create profits by minimizing costs on the one hand, or by maximizing the service level on the other hand.

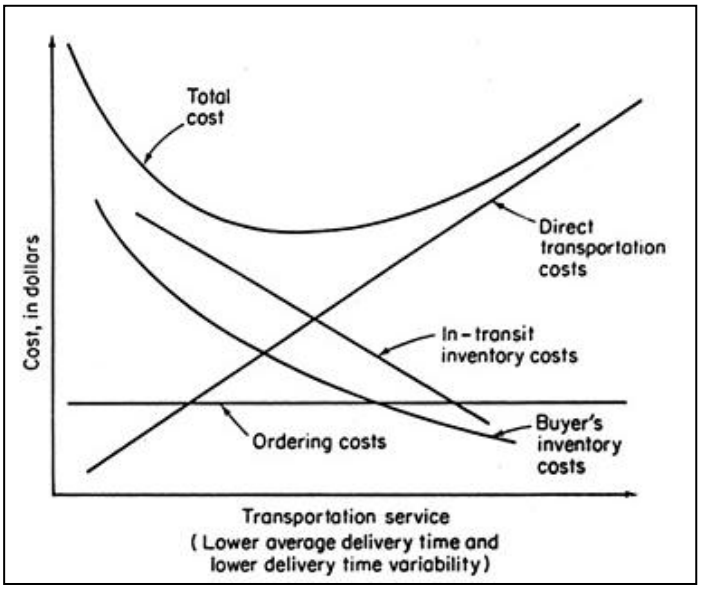


Figure 3 - Costs and Service Level

Source: U.S. Department of Transportation

Companies have to choose the right transportation solution to be competitive. Sometimes firms are only asked to provide a standard service level. In other cases, it's important for companies to improve the "customer satisfaction" because it can result in much higher perceived value (this resulting in higher prices and a better company image). In the fashion industry for example, replenishment time in the "points of sale" is a critical success factor: customers are very likely to pay more for a quick delivery and a refreshed inventory.

Firms can deliver by airplanes, ships, trucks or trains. Many times it's not easy to understand what is the best option: there are many different **"drivers"** to be considered such as product characteristics, customer requirements, countries infrastructure.

3.2 How Goods Move

*"**Transport**: take or carry (people or goods) from one place to another by means of a vehicle, aircraft, or ship."*

Source: Oxford Dictionary

There are different "modes of transportation" and each of them has different features in terms of costs, speed, reliability and accessibility. The four main transport modalities are "sea shipping", "air shipping", "road transportation" and "railway".

Each mode matches specific freight requirements so that companies should choose the most suitable transport mean.

Most of times, companies have to match different vehicles to get final destination (**intermodal transportation**). For example, it's common that a container is shipped by sea up to the harbor and then it's delivered to the final destination by truck.

Considering trans-oceanic shipping, decisions about transport modes are reduced to the choice between the air and the sea shipping alternatives (because they cover the most of the route). That's the reason why our thesis, such as the most of the economic literature, will focus on this trade-off analysis.

The overland transportation systems (road and railway) are mainly used for "local" transportation and door-to-door service.

In the following paragraph, we will analyze the different means of transportation presenting all of their technical features.

3.2.1 Sea Shipping

Sea-freight is the cheapest transportation mode; the first self-evident consideration is that ships have a cargo capacity that is considerably larger than any other transport mean (containerized vessels freight on average from 7000 to 50000 tons, while aircrafts don't outweigh the 50 tons). This allows to exploit economies of scale, and thanks to its high volume capacity, this type of transport is suitable for very bulky cargoes.

On the other hand, the sea-freight it's the slowest alternative: typically, a vessel starting from China takes 4-5 weeks to reach the US Coasts, while an aircraft requires just few days (considering all the inbound and outbound logistics activities).

As to the risk of damaging goods, the sea-freight is the least safe transport mean and the damage probability increases in case of non-containerized vessels. That's the reason why, goods that are shipped by sea require higher packaging costs. Furthermore, pilferages and losses affect sea shipment more than others means.

From a technical and infrastructural point of view, there are five different types of vessels: General Cargo, Container Ship, Tankers, Bulk & Multipurpose and RO-RO ship.

General Cargoes transport goods in bulk, in boxes and/or in pallets; the goods in a General Cargo require a resistant packaging (e.g. wooden cases) in order to assure a certain level of security.

Container Ship is an evolution of the General Cargo: in this case cargos are not loaded with different packages, but goods are disposed within containers; the containers dimensions are standardized and they are mainly available in lengths of 20 and 40 feet (see Table 2).

	20 Foot Container	40 Foot Container
Capacity (m^3)	30	60
Internal dimensions L x W x H (m)	5.89 x 2.32 x 2.23	12 x 2.32 x 2.43
Door W x H (m)	2.30 x 2.14	2.30 x 2.23
Maximum Load (tons)	18	30

Table 2 – Standard Containers

Source: UNDP

The Tanker is a special type of ship that is dedicated to the transport of liquids (e.g. petroleum derivates).

Bulk and Multipurpose vessels are dedicated to various cargo categories (both liquid and dry), that can be loaded in the same time.

Finally, the Roll on/Roll off ships (RO/RO or RORO) are special ferries dedicated to the shipment of trucks, trailer trucks and rail freight cars; this type of vessels are useful, especially for the multimodal transport.

Cargo ships can be further divided in two different categories according to the service provided: linear or tramp shipping. Liner shipment follows fixed routes, timetable and tariffs while tramp shipping is more flexible (it doesn't have fixed schedule and its price is "spot").

In the following table (table 3) technical data about the different types of vessel are summarized.

SHIP	Transport unit load	Capacity	Handling equipment
General Cargo	Dry bulk, cases, pallet loads, containers, cars	1.000 - 50.000 tons	Cranes, on board derricks
Bulk and multipurpose	Dry bulk, cases, pallet loads	10.000 - 300.000 tons	Cranes
Tankers	Liquid, gas bulk	1.000 - 500.000 tons	Pipelines
Container Ship	Containers	7.000 - 50.000 tons	Gantry cranes, straddle carriers, reach stackers
Ro-Ro Ship	Trailers, semitrailers, cars	2.500 – 20.000 tons	-

Table 3 - Ship Types

Source: Prof. A. Perego slides course "Logistics Management", 2010

3.2.2 Air Shipping

Air freight is the fastest mode of transport. According to the TNT simulator[4], the Express service takes only two working days to deliver a parcel from Milan to New York.

On the other hand, air shipment is the most costly alternative, so that it is selected to match specific requirements or constraints (such as rapid delivery times).

This transport mode is also the safest from damaging and pilferage risk. Airfreight requires less packaging than ocean and land freight.

[4] TNT simulation is available on the website:
http://www.tnt.it/TempiConsegna/intTransitTimes.do

Aircrafts can be exclusively dedicated to freight transport (e.g. the Boeing 747-400) or they can concurrently transport passengers and "pallets". In the last decades, the aviation industry has developed new types of cargo-aircraft characterized by better performances in terms of cargo capacity, speed and reliability.

Freight aircrafts have different characteristics from traditional passenger aircrafts: they have a wider body, high wings that provide a low fuselage with easy entry (and a flat cargo floor), and a higher number of wheels.

Examples of cargo airplanes are the Boing 747 and the Airbus A330 (Table 4).

	Length [m]	**Maximum Takeoff Weight [kg]**	**Maximum Cargo Volume [m³]**	**Maximum Speed [km/hr]**
Airbus A330-200	59,69	233 000	475	611
Boeing 747-400	70,6	396 890	780	901

Table 4 - Airbus 330 and Boeign 747 characteristics

Source: Boeing and Airbus

Since the cargo capacity is the principal restriction for air shipping, goods are loaded with special containers and pallets that aim at maximizing the "loading capacity"; usually pallets are covered by nets (tightened over the freight by tensioned straps), while containers are close metallic boxes whose dimensions follow specific regulations.

3.2.3 Road Shipping

Road shipping includes transport by truck, rig and wagon. This mode is the most convenient for short distances: "*for a haul that is less than 1000 km, the road freight results usually more convenient than the rail and air transport*" (UNDP source).

The main advantage of transporting by road regards the "connectivity": only by road you can go anywhere and get any final destination. That's the reason why trucks are used for "**door to door service**", playing a key role even for intermodal transportation (Road- Air-Road; Road-Sea-Road).

Road transport is satisfactory in terms of safety: thanks to the limited "handling of goods" the risk of breakage and damage is very low (in case of direct "door to door" transport, cargo is moved only at the beginning and at the end of the delivery).

Carrier companies can arrange shipping in two different ways; they can use a Full Truck Load (FTL) or a Less-than-Truck Load (LTL). The difference is mainly about the size of the transported cargo. Most of times companies use both of the modalities, to maximize saturation and minimize distance.

FTL cargo is adept at transporting large quantities to a unique destination. On the other hand, LTL trucks are dedicated to small quantities and different destinations. Carrier companies consolidate many different orders on a single FTL to minimize the distance to be done for all of the deliveries. Once the truck has reached a switching point, there's a "re-sorting" on many LTLs to get many close final destinations. Instead of dedicating a whole truck to small cargoes, companies try to exploit the available space "clustering" many different customers having small loads: risk of having costly and not-saturated shipment is removed.

3.2.4 Rail Shipping

Railway is a valid alternative to road transport and it's used for medium/short distances. Since maritime and air terminals are frequently connected to railway, the transport by train is sometimes involved in the "intermodal transportation".

Trains are less "flexible" than trucks or carriers: a train wagon must follow precise timetables and routes. Moreover, the railway network doesn't reach every area in the world and "cargo trains" are usually slower than trucks (an exception is the High Speed Train-HST that can reach 150 km/h).

3.2.5 Transport Modes Classification

The following table (Table 5) ranks the different transport modes in terms of cost, speed, punctuality and accessibility.

MODE	Cost per ton and per kilometer	Line haul transit time	Delivery time variability		Accessibility
			Absolute	%	
Road	3	2	2	1	1
Rail	2	3	3	2	2
Water	1	4	4	3	4
Air	4	1	1	4	3

Table 5 - Modes of Transport Ranking

Source: Prof. A. Perego slides course "Logistics Management", 2010

The classifications by cost and speed summarize what already said in the "technical description" while punctuality and accessibility are two points to be discussed.

As for the punctuality, aircrafts are the most reliable mean in absolute terms, while vessels are the least trustworthy one. If punctuality is considered as the percentage variation on the expected time, airplanes become the least effective mode of transport. However, this result is somehow misleading: planes have much shorter expected time and this influences the result.

The accessibility is related to the infrastructures' availability. Trucks have the highest accessibility level since the presence of roads is the only requirement to reach your destination. Regarding the other transport means, the connectivity index depends on the presence of dedicated infrastructures. Usually the more developed the countries, the better the connectivity. However, the railway network covers the territory better than airports in average, while ships are the least accessible mean of transport (only locations on the coast can be directly reached).

3.3 Transport Innovation and New Technologies

Technological improvements and costs reduction "pushed" international trade to grow. Over the years, technology has radically changed transportation, increasing both of the "quality" and the efficiency of trade.

David Hummels (2007) examines the nature of trade and the impact of technology on transportation costs. Hummels defines the latter half of the twentieth century as the "**Second Era of Globalization**": a period of rapid technological changes and costs reductions.

Even if transportation costs reduced a lot over the last years, technological development didn't impact only on cost performances. It improved speed, safety and reliability of transportation.

All of these changes influenced the "relative advantages" of the different transport systems. In the following paragraphs we go through the main technological innovations, assessing their importance during the "Second Era of Globalization", with a particular focus on air and sea transport.

3.3.1 Sea Transport Innovation

Most of the goods are traded by sea all over the world. Looking at the "Eurostat" data set, in 2010 EU27 imported 90% of the total weight and 61% of the total value of goods by sea (Figure 4). This European trend reflects the worldwide transportation tendency that the "International Chamber of Shipping" (ICS) has declared during the Ministerial Conference on Transport Congestion: "*About 90% of all world trade is transported by sea*"(2007).

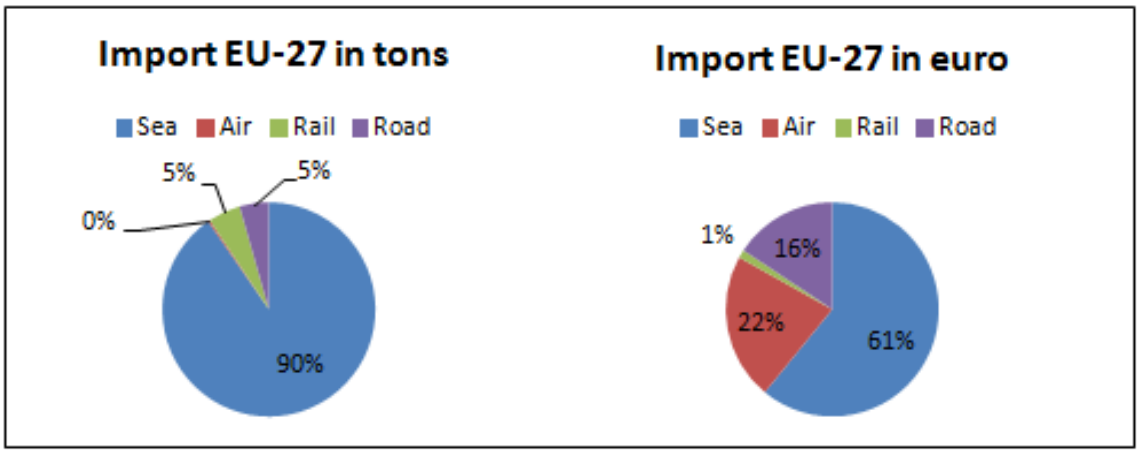


Figure 4 - European Trade by Mode of Transport

Source: Eurostat

Many technological improvements have changed the way of transporting goods by sea over the years; Hummels (2007) states that there have been three main innovations in sea transport.

The "**open registry"** allowed registering ships under "Flag of Convenience", saving time and money during berthing.

Shipping hubs like Rotterdam, Singapore and Honk Kong became more efficient because of the high-capacity vessels that lowered unit costs while increasing **transport volumes**.

The introduction of **containers** changed the way of loading, shipping and unloading cargoes. Many researchers consider the container the most important innovation in transportation after the World War II.

In 2009, the economist Paul Krugman writes: *"When we think about technology that changed the world, we think about glamorous things like the internet. But if you try to figure out what happened to world trade, there is a really strong case to be made that it was the container".*

"Containerization" was introduced in the US during the '60s and it spread worldwide during the '70s, "improving trade" in many ways.

Economists show different points of view about the effects of containerization. Hummels (2007) analyzes the situation from a financial perspective. He points out that containers didn't bring radical improvements because shipping costs lowered only from 3 to 13 percent in average.

Looking at data from UNCTAD, (Figure 5: the available fleets in the world, divided by type of ship) it's clear that containerization is spreading fast: in the last three decades the quantity of container ships has increased more than fifteen times and most of the countries have converted their harbour facilities to manage containers.

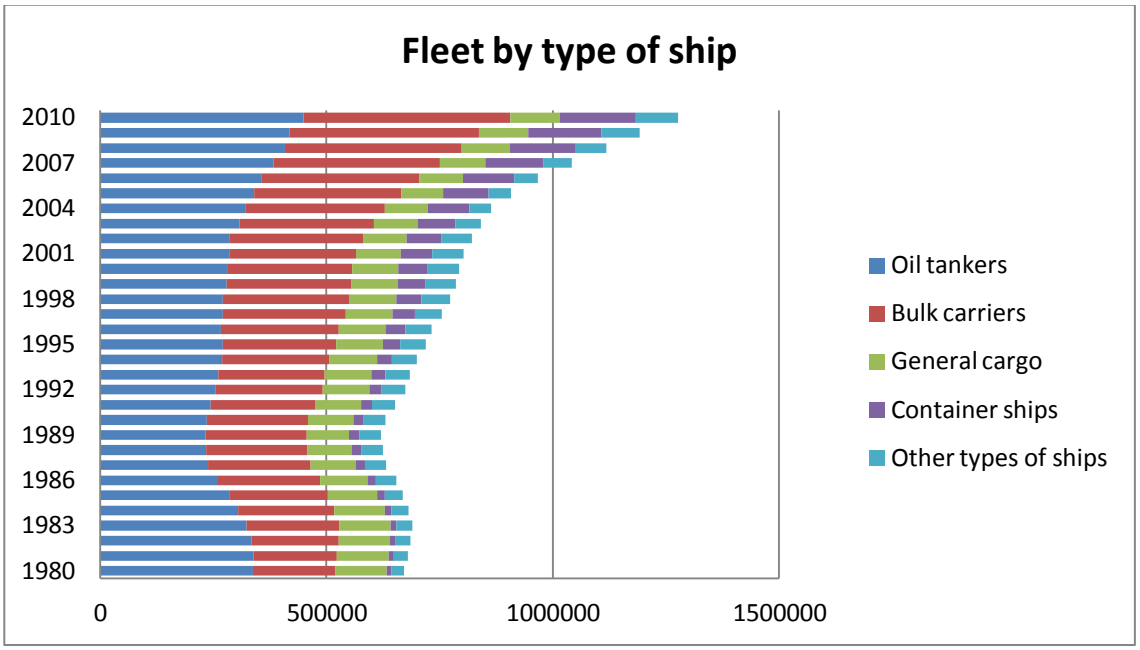


Figure 5 - Worldwide Fleer by Type of Ship

Source: UNCTAD

In a recent analysis, Bernhofen et al. (2011) show that there is a positive correlation between the introduction of containers and international shipping. They demonstrate through a gravity model that bilateral trade went up by 23% after containerization.

Rather than reduction of costs, containerization brought also many other advantages: time saved in uploading/downloading, the possibility to fasten the intermodal transport, the opportunity to exploit economies of scales with a better space utilization and the decrease of pilferage (through the containers' protection).

3.3.2 Air Transport Innovation

New technologies have changed “air transport” during the second half of the twentieth century. Improvements to **aerodynamics**, new innovative **materials** and the adoption of **jet engines** made air freight more reliable, faster, less needy of maintenance and cheaper. In the same period, companies started building airplanes only dedicated to the transport of cargo.

Hummels (2007) analyzes “air-freight prices” from 1955 to 2004. Prices shrank after the “jet introduction” and they kept going down in the three decades thereafter. Airplanes became much more convenient than before because the average cost difference between air freight and sea shipment was strongly reduced.

If this trend won’t change in future, air transport is going to take a more and more important role in trade.

3.4 International Agreements and Air Transport Costs

In the last two decades, many international agreements radically modified the competitive context in air transport industry. Between 1992 and 2003, the United States signed a series of **Open Skies Agreements** bringing a radical change in air trade equilibrium. New opportunities were born and the **air trade deregulation process** contributed to the reduction of air transportation costs and to the consequent attenuation of "fares gap" between air and sea shipping modes.

In 1992, the United States signed the first Open Skies Agreement with Netherland. In the following years, the United States entered into Open Skies Agreements with over 100 partners from every region of the world and at every level of economic development.

The primary objectives of OSAs are:

- To liberalize the aviation market eliminating government interference in the commercial decisions about routes, capacity, and pricing, and freeing carriers to provide more affordable, convenient, and efficient air service for consumers[5].

- To stimulate and promote international air trade by allowing the airline companies to offer the travelling and shipping public a variety of service options at the lowest price (A.Micco, T.Serebrisky, 2006).

The OSAs were implemented both in passengers and cargo air transportation.

[5] Source: U.S. Department of State

The key Open Skies provisions[6] are:

1. ***Free Market Competition:*** *No restrictions on international route rights; number of designated airlines; capacity; frequencies; or types of aircraft.*
2. ***Pricing Determined by Market Forces:*** *A fare can be disallowed only if both governments concur -- "double-disapproval pricing" -- and only for certain, specified reasons intended to ensure competition.*
3. ***Doing Business Protections:***
   - *All carriers of both countries may establish sales offices in the other country, and convert earnings and remit them in hard currency promptly and without restrictions.*
   - *Carriers are free to provide their own ground-handling services -- "self handling" -- or choose among competing providers. Airlines and cargo consolidators may arrange ground transport of air cargo and are guaranteed access to customs services.*
   - *User charges are non-discriminatory and based on costs.*
4. ***Cooperative Marketing Arrangements:*** *Airlines may enter into code-sharing or leasing arrangements with airlines of either country, or with those of third countries. An optional provision authorizes code-sharing between airlines and surface transportation companies.*
5. ***Provisions for Consultation and Arbitration:*** *Model text includes procedures for resolving differences that arise under the agreement.*
6. ***Liberal Charter Arrangements:*** *Carriers may choose to operate under the charter regulations of either country.*

---

[6] Open Skies provisions are reported by the "U.S. Department of State": http://www.state.gov/e/eeb/rls/fs/2009/119760.htm

7. ***Safety and Security:*** *Each government agrees to observe high standards of aviation safety and security, and to render assistance to the other in certain circumstances.*
8. ***Optional 7th Freedom All-Cargo Rights:*** *Provides authority for an airline of one country to operate all-cargo services between the other country and a third country, via flights that are not linked to its homeland.*

A. Micco and T.Serebrisky (2006) demonstrate a positive correlation between OSA's introduction and air shipping costs decrease. In particular, they find that *OSAs reduced air transport costs by about 9% and increased by 7% the share of imports arriving by air within 3 years after an OSA is signed.* They discovered also that effects of OSAs are not homogeneous in all the countries. For low-income developing countries the OSAs' effects are very restricted; they interpreted this result as an indication that developing countries were not able to exploit the OSAs' advantages because of their limits in infrastructures and their "small relevance" in international trade.

3.5 Transport Fares: Empirical Evidence

Air shipping is much more expensive than sea freight. But, what is the numerical dimension of transportation costs? What is the real difference between fares of different transport modes?

In this chapter we want to determine fares' order of magnitude for the different shipping modes. We have collected numerical data from the "fares calculator systems" of two American logistic companies: the FedEx Corporation[7] for air shipping and the IQ Global Logistic Corporation[8] for sea freighting.

The web applications permit to implement detailed queries by inputting data about origin and destination countries and product characteristics (dimensions and weight).

We made a first query for the New York-Genoa route, considering a 50cm x 50 cm x 50cm (Length x Width x Height) pack. The output we obtained is 955 € for air delivery and 143€ for sea shipping; in this case air transportation rate is 568% bigger than sea shipping one.

However, a single observation cannot represent a general rule (we cannot state that air fare is always 568% bigger than the sea one). Many other queries are necessary to obtain a valid order of magnitude for costs' differential.

We know from theory that many factors (e.g. distance, volume, weight…) affect transportation costs. These determinants influence transport rates in different ways according to different means. We decided to analyze the transportation rates by varying factors one at a time and keeping constant all the other factors.

[7] FedEX: https://www.fedex.com/ratefinder/home?cc=US&language=en&locId=express

[8] IQ Global: http://www.iqglobal.com/

3.5.1 Distance

Transportation costs vary according to **distance**. Our goals are:

- To understand in which way distance influences the different transportation fares.
- To analyze the slacks between air and sea fares in relation to distance.

The first consideration we have to make is about the kilometric difference between the "air and sea" route. If a ship goes from Genoa to Palermo, the kilometers to be travelled are close to those travelled by plane, but it's not always like this (table 6). For instance, considering the route from Livorno to New York, the maritime distance is 7572 km (4086 nautical miles), while the air distance is only 6638 km. The "distance gap" is even higher for other routes such as Rotterdam-Melbourne or Genoa-Tokyo.

	Nautical Distance	**Air Distance**	**%Δ (sea - air)**
Genoa/Palermo	795 km (429 n.m.)	792 km	+ 0,4%
Livorno/ New York	7572 km (4086 n. m.)	6638 km	+ 14,07%
Rotterdam/ Melbourne	20544 km (11086 n. m.)	16571 km	+ 23,98%
Genoa/ Tokyo	17303 km (9337 n. m.)	9828 km	+ 76,06%

Table 6 - Nautical Distance and Air Distance
Source: sea.rate.com simulator

Here the maps with the nautical routes (the route to go in Tokyo almost doubles by ship).

Livorno – New York (+14% of Km)



Figure 6 - "Sea" and "Air" Routes from Livorno to New York
Source: sea.rate.com simulator

Rotterdam – Melbourne (+24% of Km)



Figure 7 - "Sea" and "Air" Routes from Rotterdam to Melbourne
Source: sea.rate.com simulator

Genoa – Tokyo (+76% ok Km)



Figure 8 - "Sea" and "Air" Routes from Genoa to Tokyo
Source: sea.rate.com simulator

Air and nautical distances can be considerably different; this is first evidence that cost by air and ship can't vary in the same way according to real routes (a difference in kilometric hauls is also a difference in transportation costs).

We made queries in the two web applications keeping constant the following parameters:

- Pieces number: 1
- Pack dimensions: 50cm x 50 cm x 50cm (Length x Width x Height)
- Pack weight: 50 kg
- Value declared (only for air shipment): 100€
- Type of good (only for air shipment): "Clothing category"

We changed the destination country in any query, aiming at obtaining observations for different distances; the origin country is always New York.

In the following table (Table 7) we report results:

Route --> New York to:	Air fare	Sea fare	Air distance [km]	Sea distance [km]	Δdistance (sea-air) [km]	Δ%distance (sea-air)
Kingston	€ 236	€ 210	2536	2709	173	6,82%
Santa Marta/ Barranquilla	€ 522	€ 309	3310	3298	-12	-0,36%
Veracruz	€ 315	€ 262	3191	3609	418	13,10%
Caracas	€ 241	€ 182	3436	3576	140	4,07%
Panama	€ 527	€ 220	3644	3737	93	2,55%
Lisboa	€ 919	€ 188	5429	5459	30	0,55%
Liverpool	€ 512	€ 171	5330	5515	185	3,47%
Bergen	€ 755	€ 241	5618	5880	262	4,66%
Genoa	€ 955	€ 143	6508	7458	950	14,60%
Rio de Janeiro	€ 700	€ 188	7766	8769	1003	12,92%
Izmir	€ 1.254	€ 186	8137	8917	780	9,59%
Tokyo	€ 1.470	€ 168	10860	17977	7117	65,53%
Melbourne	€ 1.116	€ 181	16690	18379	1689	10,12%
Beijing	€ 1.004	€ 306	11001	19155	8154	74,12%
Bangkok	€ 1.065	€ 180	13950	20209	6259	44,87%

Table 7 – Fares and Distance analysis
Sources: FedEx, IQ Global and sea.rate.com

Considering air shipping, fare has a positive correlation with distance. Looking at "Figure 9", it's clear that fares grow as a "logarithmic function".

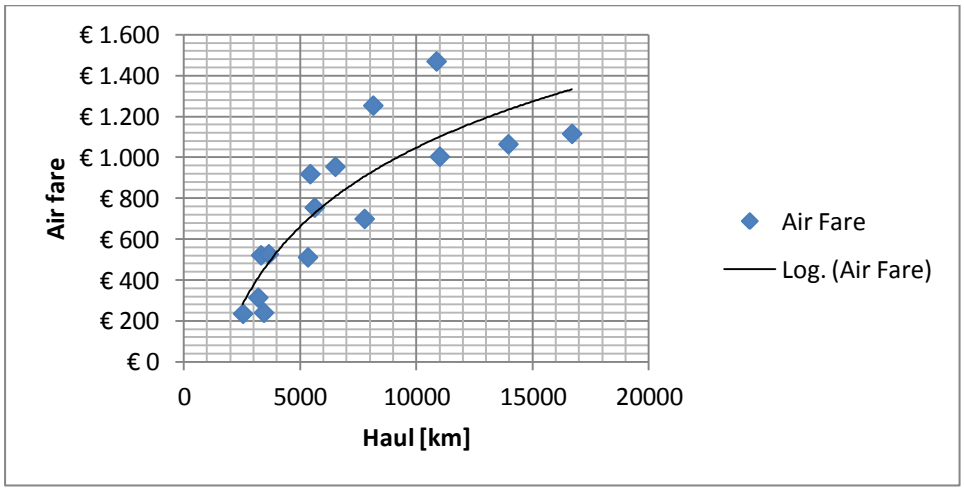


Figure 9 - Air Shipping Fare depending on Distance
Source: FedEx

The same relationship is not confirmed in case of sea shipment. In particular, **the sea fare seems to be unrelated to distance** (Figure 10).

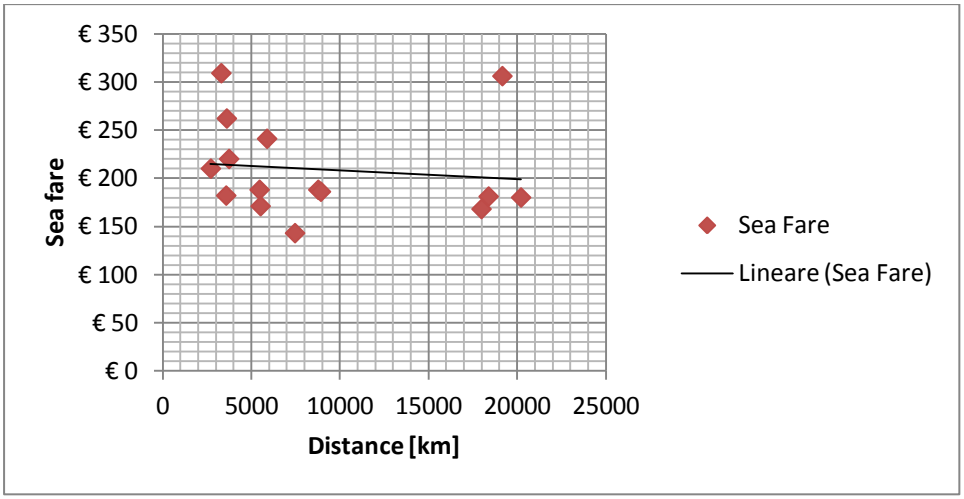


Figure 10 - Sea Shipping Fare depending on Distance
Source: IQ Global

A reasonable explanation for these two opposite results is the difference between air and sea variable costs. In general, the longer the distance the higher the variable costs (e.g. fuel cost); however, **sea shipping variable costs are considerably lower than air shipping ones** (in table 8 we report the average transport costs). Sea Freight could be more affected by route types and infrastructures availability than by distance.

Transport Mode	c€/(ton*km)
Rail	1,25
Road	12,5
Sea	0,35
Air	29,00

Table 8 - Variable Costs by Mode of Transport
Source: Valocchi

According to queries' results, the gap between the two fares (air and sea) is very variable (Table 9).

Route --> New York to:	Air fare	Sea fare	Δ fare (air-sea)	Δ% fare (air - sea)	Air distance [km]	Sea distance [km]
Kingston	€ 236	€ 210	€ 26	12%	2536	2709
Santa Marta/ Barranquilla	€ 522	€ 309	€ 213	69%	3310	3298
Veracruz	€ 315	€ 262	€ 53	20%	3191	3609
Caracas	€ 241	€ 182	€ 59	32%	3436	3576
Panama	€ 527	€ 220	€ 307	140%	3644	3737
Lisboa	€ 919	€ 188	€ 731	389%	5429	5459
Liverpool	€ 512	€ 171	€ 341	199%	5330	5515
Bergen	€ 755	€ 241	€ 514	213%	5618	5880
Genoa	€ 955	€ 143	€ 812	568%	6508	7458
Rio de Janeiro	€ 700	€ 188	€ 512	272%	7766	8769
Izmir	€ 1.254	€ 186	€ 1.068	574%	8137	8917
Tokyo	€ 1.470	€ 168	€ 1.302	775%	10860	17977
Melbourne	€ 1.116	€ 181	€ 935	517%	16690	18379
Beijing	€ 1.004	€ 306	€ 698	228%	11001	19155
Bangkok	€ 1.065	€ 180	€ 885	492%	13950	20209

Table 9 - Fares Difference

Sources: FedEx, IQ Global and sea.rate.com

On the one hand, the fare differential is limited between 12% and 69% for short-hauls: if distance to be covered is short, variable costs are less relevant and fixed costs (high for both of the means) are the most important cost component.

On the other hand, the gap is significantly larger for long distances (in the New York – Genoa route the gap exceeds the 500%).

Let's try to represent the price differentials depending on distance (Figure 11):

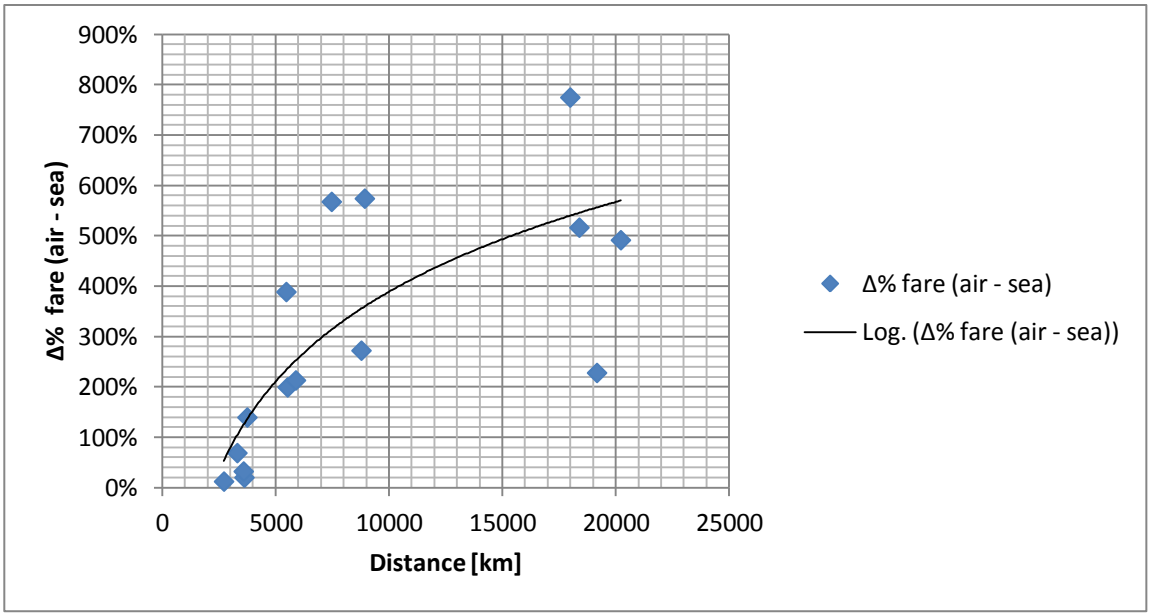


Figure 11 - Fares Difference depending on Distance

Sources: FedEx, IQ Global

The gap between air and sea fares increases with distance. This result confirms that variable costs are very different for the two transportation means.

In some cases, the fares differential is influenced also by the kilometric difference between air and sea routes. For example, the route New York -Tokyo has the highest fare gap (Δ% fares= 775%) and the kilometric difference between the two transport mode is about 60% (10860 km for the air route and 17977 km for the sea path).

3.5.2 Weight and Volume

Volume and weight impact on transportation costs. We try to determine how these cargo characteristics influence fares.

Aircrafts are more sensitive to weight and volume because of the limited cargo capacity. Air prices are usually set by making a comparison between the "Actual Weight" and the "Dimensional Weight" (where "Dimensional Weight" is a derived weight that takes into account the volume occupied by cargo: further details at 4.3.1 "value/weight"); if the "Dimensional Weight" is higher than the "Actual Weight", the first one is used to set the fare.

On the contrary, sea shipping has not rigid boundaries in terms of volume and weight.

Weight

Weight is a relevant constraint to deliver by air because higher weights imply stronger extra-costs. On the contrary, we can make the assumption that weight is not very relevant in sea-shipping. Indeed, the company taken into consideration don't consider weight as a pricing factor until the cargo to be delivered don't exceed a certain given density limit (750 kg per cubic meter).

Focusing on air shipping, we made several queries fixing these parameters:

- Origin: New York
- Destination: Milan
- Pieces number: 1
- Pack dimensions: 50cm x 50 cm x 50cm (Length x Width x Height)
- Value declared: 100€
- Type of good: "Clothing category"

In the following table and graph we report the queries' results.

Weight [kg]	Air Fare
1	€ 523
10	€ 523
20	€ 523
30	€ 610
50	€ 907
70	€ 1.200
80	€ 1.387
100	€ 1.699
200	€ 3.398
300	€ 4.779
400	€ 6.372
500	€ 7.729
600	€ 9.275
700	€ 10.821
800	€ 12.366
900	€ 13.912

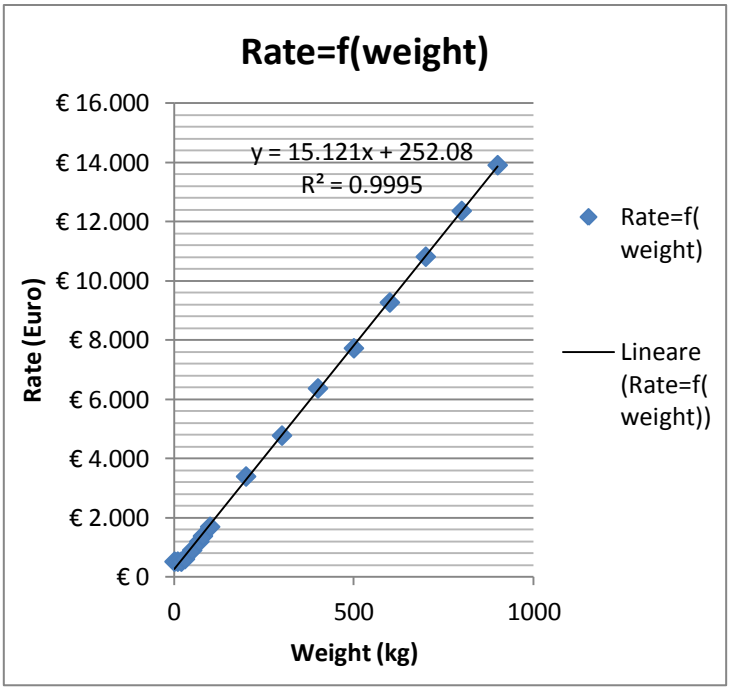


Figure 12 - Air Fares depending on Weight

Table 10 - Air Fares depending on Distance

Source: FedEx

From data, we can conclude that air fare increases almost linearly.

Volume

We made queries for both air and the sea shipment in order to understand how fares vary according to cargo volume (Table 11). Here the fixed parameters:

- Origin: New York
- Destination: Livorno
- Pieces number: 1
- Pack weight: 50 Kg
- Value declared (only for air shipment): 100€
- Type of good (only for air shipment): "Clothing category"

Length [cm]	Width [cm]	Height [cm]	Volume [m^3]	Weight [kg]	Dimensional Weight [kg]	Sea Fare	Air Fare
10	10	10	0,001	50	0,2	€ 177	€ 908
50	50	50	0,125	50	25,0	€ 177	€ 908
60	60	60	0,216	50	43,2	€ 177	€ 908
62,99	62,99	62,99	0,250	**50**	**50,0**	€ 177	€ 908
70	70	70	0,343	50	68,6	€ 177	€ 1.178
80	80	80	0,512	50	102,4	€ 177	€ 1.733
90	90	90	0,729	50	145,8	€ 177	€ 2.480
100	100	100	1,000	50	200,0	€ 177	€ 3.398
115	115	115	1,521	50	304,2	€ 191	€ 4.827
126	126	126	2,000	50	400,1	€ 235	€ 6.356
136	136	136	2,515	50	503,1	€ 288	€ 7.744
150	150	150	3,375	50	675,0	€ 371	€ 10.387
160	160	160	4,096	50	819,2	€ 443	€ 12.613
170	170	170	4,913	50	982,6	€ 526	€ 14.986
175	175	175	5,359	50	1071,9	€ 570	€ 15.990
200	200	200	8,000	50	1600,0	€ 790	n.a.
220	220	220	10,648	50	2129,6	€ 1.008	n.a.
250	250	250	15,625	50	3125,0	€ 1.474	n.a.
270	270	270	19,683	50	3936,6	€ 1.855	n.a.
300	300	300	27,000	50	5400,0	€ 2.546	n.a.

Table 11 - Fares depending on Volume

Source: FedEx, IQ Global

In both of cases, there is a minimum charge value (177 € for sea shipping and 908€ for air freight). Air shipping presents also a maximum chargeable volume because Fedex Corporation doesn't provide any service over this limit. Air fare is calculated by comparing "actual weight" and "dimensional weight" (DW= (L x W x H)/5000); in this case, "dimensional weight" exceeds "actual weight" if volume is higher than 0,25 cubic meters (we used a fixed weight in queries).

From data we confirm that both of air and sea fares are positively correlated with volume. But, does volume influence fares in the same way? Looking at the graph (Figure 13: relationship between fare and volume) we observe that "air fare curve" is steeper than "sea fare curve". Volume is a more "influent" factor in air pricing (air fare increases more than sea fare if volume increases by one unit).

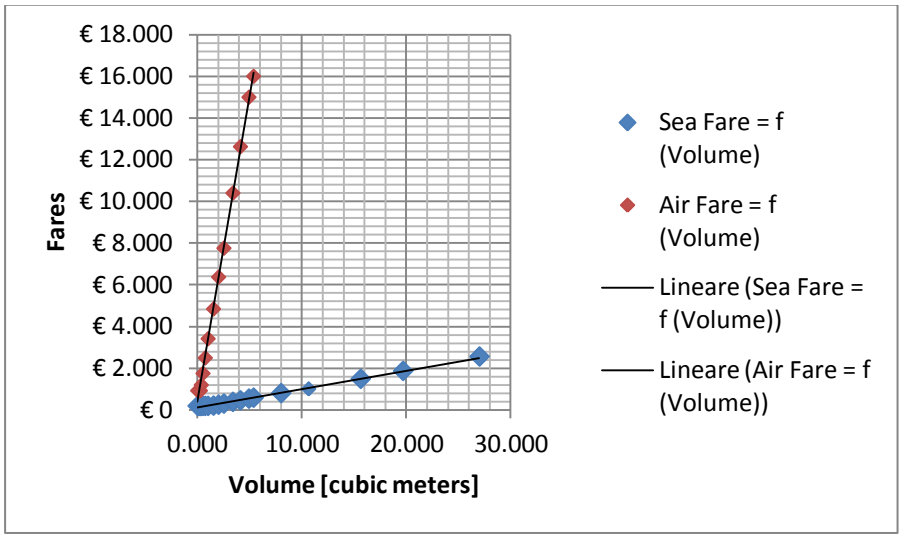


Figure 13 - Fare and Volume Relationship

Source: FedEx, IQ Global

In the following table (Table 12), we report fare percentage slacks, according to different volumes of cargo. The gap between air and sea fare is higher if volume is higher: the bulkier the cargo, the more convenient the sea transport mode.

Volume [m^3]	Sea Fare	Air Fare	Δ Fare %	Volume [m^3]	Sea Fare	Air Fare	Δ Fare %
0,001	€ 177	€ 908	412,50%	**2,515**	€ 288	€ 7.744	2592,40%
0,125	€ 177	€ 908	412,50%	**3,375**	€ 371	€ 10.387	2702,37%
0,216	€ 177	€ 908	412,50%	**4,096**	€ 443	€ 12.613	2745,27%
0,250	€ 177	€ 908	412,50%	**4,913**	€ 526	€ 14.986	2751,32%
0,343	€ 177	€ 1.178	564,90%	**5,359**	€ 570	€ 15.990	2704,97%
0,512	€ 177	€ 1.733	878,15%	**8,000**	€ 790	n.a.	
0,729	€ 177	€ 2.480	1299,78%	**10,648**	€ 1.008	n.a.	
1,000	€ 177	€ 3.398	1817,92%	**15,625**	€ 1.474	n.a.	
1,521	€ 191	€ 4.827	2433,67%	**19,683**	€ 1.855	n.a.	
2,000	€ 235	€ 6.356	2604,77%	**27**	€ 2.546	n.a.	

Table 12 – Fares Difference according to Volume

Source: FedEx, IQ Global

Finally, we represent data in a graph (Figure 14). We obtain an "S shape" curve in which we can recognize three homogeneous areas:

1) Gap is constant for small volumes (due to the minimum charge value imposed for both air and sea shipping).
2) The curve is steeper for volumes between 0,25 and 2 cubic meters: a volume variation increases considerably the fares gap.
3) The differential has lower slope again for volumes over 2 cubic meters.

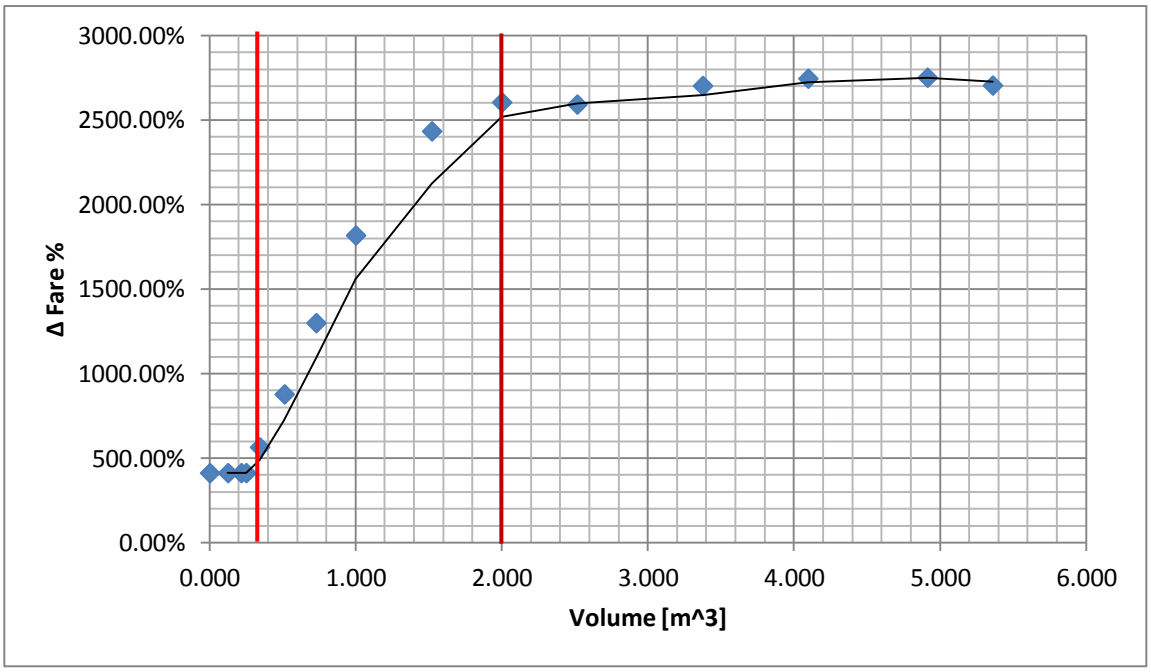


Figure 14 - Fares Difference according to Volume

Source: FedEx, IQ Global

3.6 Focus on European Trade

Before entering in detailed data analysis, we carried out a preliminary investigation about the European trade.

Which modes of transportation does the European Union use? What does the European Union trade? Which are the main trade partners of the European Union? These are some of questions we want to answer in this preliminary analysis.

Data Base Structure

Firstly we determined the "geographical boundaries" of our analysis. The European Union (27 countries) is the reference point. Then we selected eight partner countries: China, India, Japan, the United States, Australia, Brazil, Canada and Russia. The selecting criteria that we followed are:

- Country relevance in international trade (see Table 13).
- Geographical position of countries: we try to cover different distances from EU27, including both countries with a direct land connection with the European Union (e.g. China and Russia) and countries with sea connection only (e.g. Australia, Brazil).

We also considered data about all of the "EXTRA-EU27" countries together: this query gives an aggregated view on European import from all over the world.

For now, we decided to focus on European imports (rather than exports) in order to understand in which categories the other countries are specialized; in this way we avoid the risk to leave out some product categories that The EU27 doesn't export (for example the EU doesn't have many minerals' reserves).

Rank	Exporters	Value	Share	Annual % change	Rank	Importers	Value	Share	Annual % change
1	Extra-EU (27) exports	1525	16.2	-21	1	Extra-EU (27) imports	1672	17.4	-27
2	China	1202	12.8	-16	2	United States	1604	16.7	-26
3	United States	1057	11.2	-18	3	China	1006	10.5	-11
4	Japan	581	6.2	-26	4	Japan	551	5.7	-28
5	Korea, Republic of	364	3.9	-14	5	Hong Kong, China	353	3.7	-10
						- retained imports a	91	0.9	-8
6	Hong Kong, China	330	3.5	-11	6	Canada	330	3.4	-21
	- domestic exports a	15	0.2	-9					
	- re-exports a	314	3.3	-11					
7	Canada	316	3.4	-31	7	Korea, Republic of	323	3.4	-26
8	Russian Federation	304	3.2	-36	8	Singapore	246	2.6	-23
						- retained imports b	114	1.2	-28
9	Singapore	270	2.9	-20	9	India	244	2.5	-24
	- domestic exports	138	1.5	-21					
	- re-exports	132	1.4	-19					
10	Mexico	230	2.4	-21	10	Mexico	242	2.5	-24
11	Taipei, Chinese	204	2.2	-20	11	Russian Federation c	192	2.0	-34
12	Saudi Arabia a	189	2.0	-40	12	Taipei, Chinese	175	1.8	-27
13	United Arab Emirates a	175	1.9	-27	13	Australia	165	1.7	-17
14	Switzerland	173	1.8	-14	14	Switzerland	156	1.6	-15
15	Malaysia	157	1.7	-21	15	Turkey	141	1.5	-30
16	India	155	1.6	-20	16	United Arab Emirates a	140	1.5	-21
17	Australia	154	1.6	-18	17	Thailand	134	1.4	-25
18	Brazil	153	1.6	-23	18	Brazil	134	1.4	-27
19	Thailand	152	1.6	-14	19	Malaysia	124	1.3	-21
20	Norway	121	1.3	-30	20	Saudi Arabia a	92	1.0	-20
21	Indonesia	120	1.3	-14	21	Indonesia	92	1.0	-28
22	Turkey	102	1.1	-23	22	South Africa a	72	0.7	-28

Table 13 - Countries Ranking by Export/Import

Source: WTO – 2010 Press Releases

We collected data from the Eurostat database: European imports are recorded both in value (Euro) and in weight (tons). Data[9] are subdivided by mode of transport and by type of goods according to the NSTR goods classification.

[9] Eurostat web site: Section "External Trade" → External trade detailed data → EXTRA EU27 – Trade Since 1999 by Mode of Transport (NSTR)

Eurostat classifies the mode of transport according to the mean that crosses the destination country's boundaries[10] so it's possible that an Australian export towards Europe is recorded under "Road" transport mean (even if Australia doesn't have a direct land connection with Europe!).

For this preliminary overview we referred to 2010 (the most recent available period).

Data Overview

Which modes of transportation does Europe use?

European Union exploits all the available transportation means for trading. Ships and aircrafts are the most common transport modes: EU27 transports 61% of total value by sea, 22% by air, 16% by road and 1% by rail (Figure 15: Extra-EU27).

For all of the eight country-partners, the percentage "by sea" (in value) exceeds 50% (with the exception of the United States). Brazilian datum is prominent: Brazil trades over 90% of the total value by sea (we will discover that the average value/weight of Brazilian goods is very low).

Rail is the least common transport mean: "rail percentage" doesn't exceed 1% of the total value, excepting for Russia and China (both countries have a direct "land connection" with EU27).

[10] Eurostat Nota Bene: The mode of transport corresponds to the active mode of transport at the entry or exit to/from the borders of the European Union. For this reason the mode of transport do not always correspond to the national border's characteristics. For example we can have flows by sea mode for declarant Austria meaning that the commodities came by sea at EU borders and the custom clearance was done in Austria.

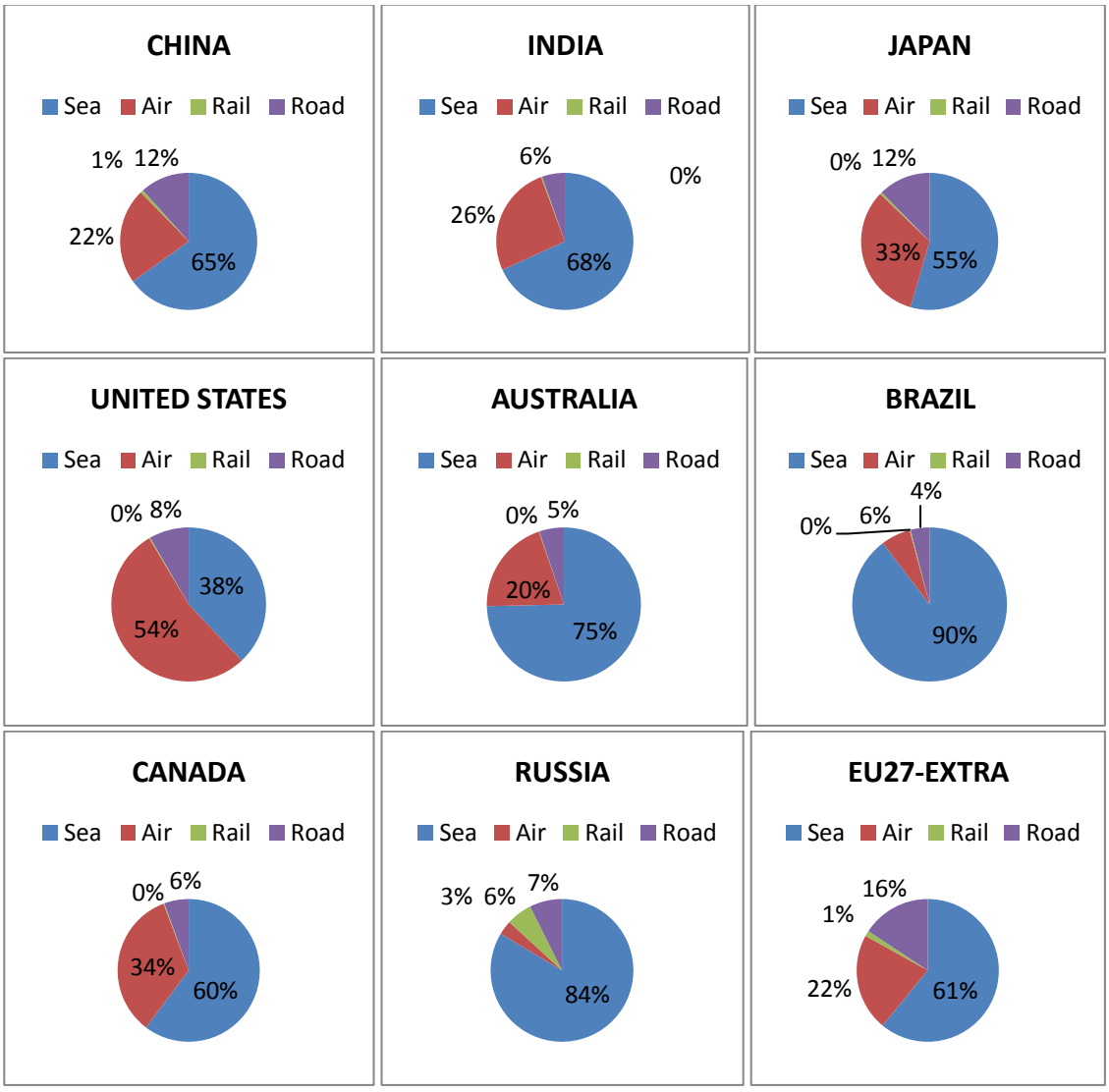


Figure 15 - Countries Trade by Mode of Transport

Source: Eurostat

Which are the main trade partners of Europe?

Let's try to compare the different countries.

In value terms, the European Union has imported more from China. Then, there are the United States, Russia, Japan, India and Australia (Figure 16).

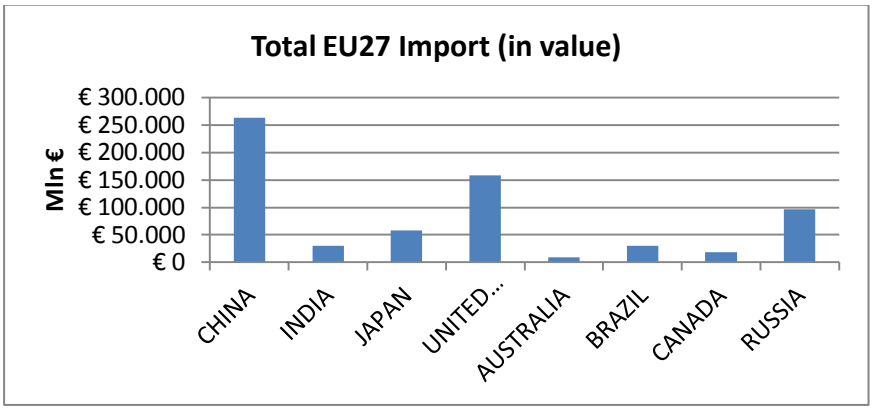


Figure 16 - Total European Import (in value) from Different Countries

Source: Eurostat

Similarly, we made this type of analysis subdividing the value imported by mode of transport (Figure 17).

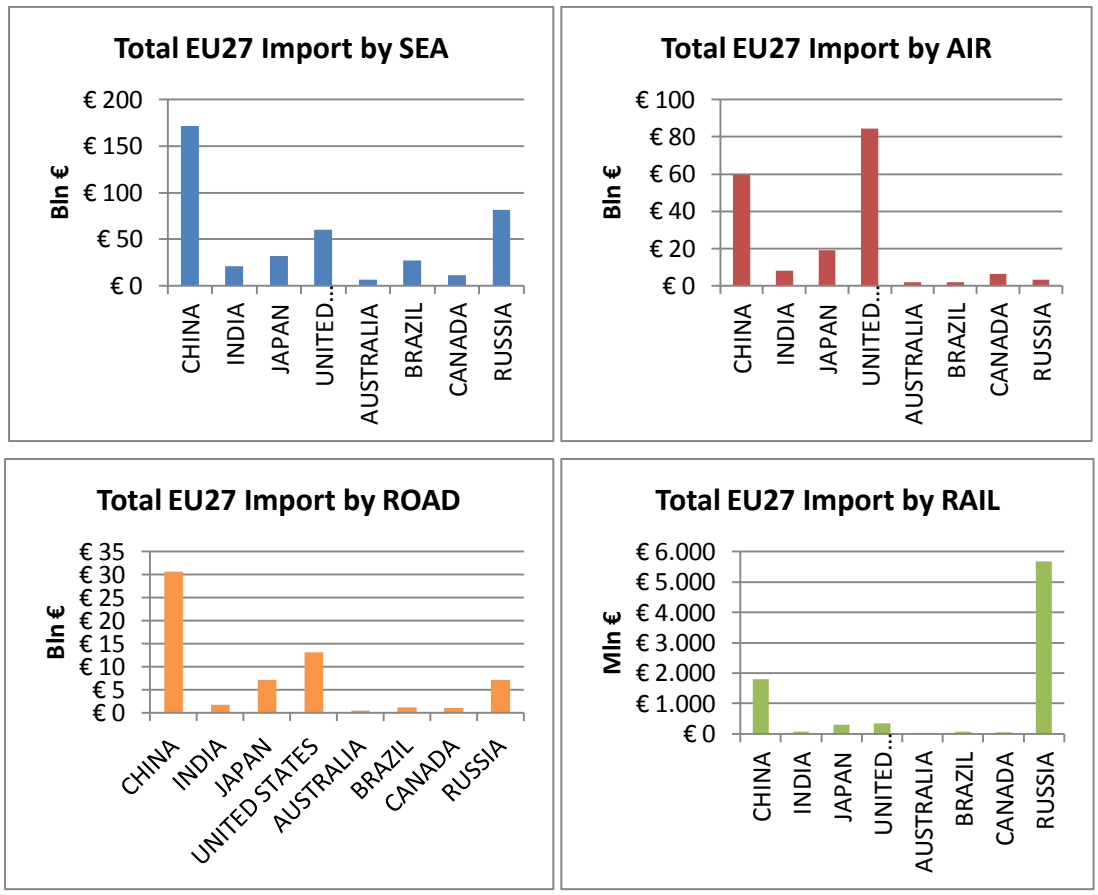


Figure 17 - Total European Import (in value) from Different Countries by Mode of Transport

Source: Eurostat

Only “road transport” (figure 17) reflects proportions identified by the “overall trade graph” (figure 16).

The United States export “by sea” less than Russia, but they are the first exporter country “by air”. We could suppose that the United States are specialized in high value-weight goods (and for those goods they select the air transport choice).

Russia has the highest rail transportation value. We could suppose that Russia can transport more by rail because it has a direct connection with the European boundaries and a more developed railway network.

Analyzing the same data under the weight perspective (in tons), we can make some considerations (we reported graph in the Appendix B):

- The most of “weight imported” by EU27 is transported by sea (over 90%).
- 11% of the total weight imported from Russia crosses the European boundaries by train (on the contrary, other countries don’t exceed 1% by rail).

Comparing the total value with the overall weight imported by EU27 from the different countries (Figure 18), we noted that Russia is the major exporter-country for EU27 in terms of weight, while it is only the third in terms of value. We can affirm that Russia exports low value/weight goods. We noted similar scenario for Brazilian and Australian data.

On the contrary China, Japan and the United States have a higher value in Euros than in kilos; these countries export goods with high value weight (e.g. High-Tech sector).

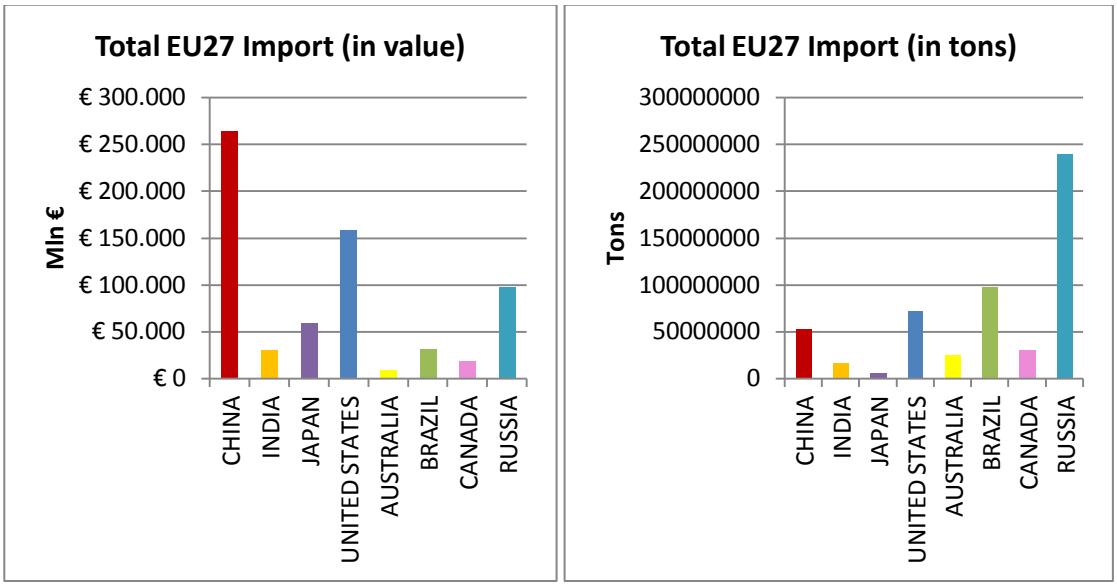


Figure 18 - Comparison between Total Value and Total Weight of European Import

Source: Eurostat

Now let's focus on value/weight comparing the different transport means (Figure 19).

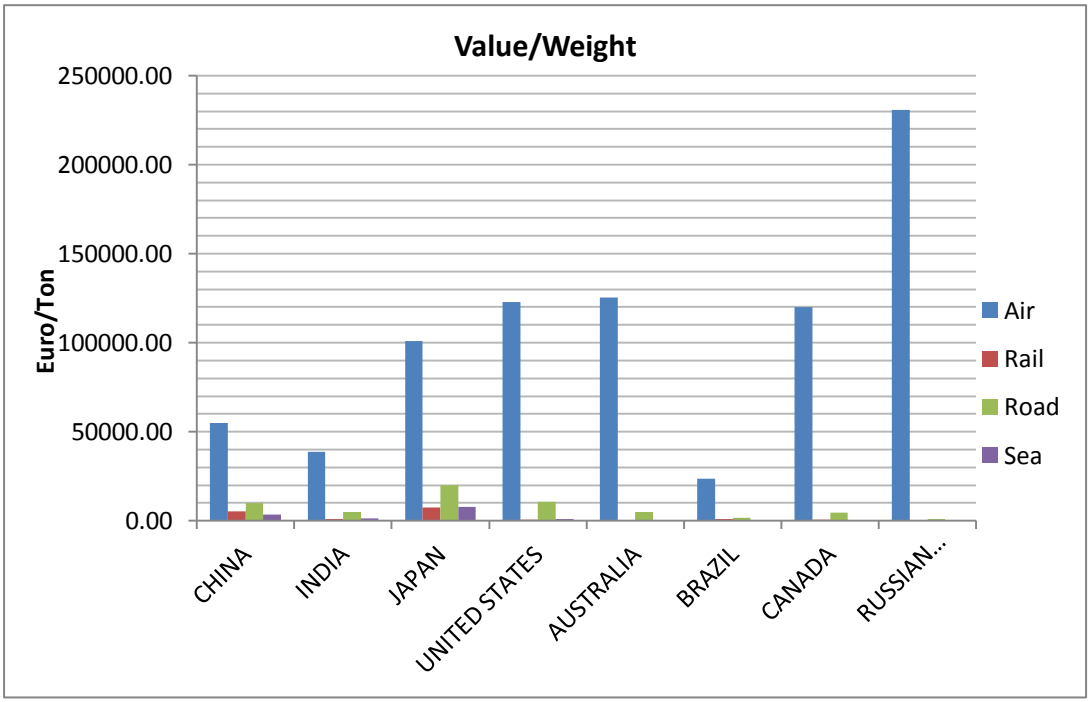


Figure 19 - Value/Weight Comparison

Source: Eurostat

From data we can observe that the "air shipping" value/weight is much higher than ratios related to other means. Goods transported "by road" have higher value/weight than those goods transported by sea and rail modes (but lower than "air transport").

The highest value/weight in dataset corresponds to Russian export by "air". Russia borders European Union and the most relevant cities like Moscow and St. Petersburg are relatively close to the European boundaries; for this reasons, "air shipment" is only 3% of total import from this country: air transport costs are too high for this short distance (excepting for very valuable goods). China, India and Brazil have lower value/weight in "air transport" on average. Long distances could oblige the air transportation option also for goods with lower value weight (due to strict time constraints). However, distance can't be the only influencing factors. For example, we should also consider the types of goods in which countries are specialized: if a country is more "labour-intensive" than "capital-intensive", its products shouldn't have high value/weight.

In Appendix B, we report all the graphs about "value/weight analysis" for each country.

What does the European Union trade?

In 2010, the European Union had a Current Account deficit of 152 billion of Euros. The main product categories **imported** by EU27 are (Table 14):

93 - OTHER MACHINERY APPARATUS AND APPLIANCES…	22,63%
31 - CRUDE PETROLEUM	15,21%
97 - OTHER MANUFACTURED ARTICLES	9,08%
96 - LEATHER, TEXTILES AND CLOTHING	7,33%
91 - TRANSPORT EQUIPMENT	6,55%
89 - OTHER CHEMICAL PRODUCTS	5,75%
99 - MISCELLANEOUS ARTICLES	4,38%
32 - FUEL DERIVATIVES	3,53%

Table 14 - Main Product Categories Imported by EU27

Source: Eurostat

The main product categories **exported** by EU27 are (Table 15):

93 - OTHER MACHINERY APPARATUS AND APPLIANCES…	28,34%
91 - TRANSPORT EQUIPMENT	13,62%
89 - OTHER CHEMICAL PRODUCTS	12,87%
97 - OTHER MANUFACTURED ARTICLES	10,07%
99 - MISCELLANEOUS ARTICLES	4,74%
32 - FUEL DERIVATIVES	4,10%
81 - BASIC CHEMICALS	3,66%
96 - LEATHER, TEXTILES AND CLOTHING	3,43%

Table 15 - Main Product Categories Exported by EU27

Source: Eurostat

It's evident that the most traded categories are almost the same for import and export (excepting for crude petroleum: Europe doesn't have a disposal of natural petroleum!). This is the effects of the **intra-trade** (trade within the same industry between similar countries) that has grown largely in the last half a century; a reasonable explanation for this intra-trade's growth is the combined effect of **product varieties' increase** (same sector but very differentiated outputs) and **economies of scale**. Paul Krugman (1995) justifies intra-trade's increase by stating that: *"…manufactured goods today are more complex than those of our great-grandfathers' day; not only are they more finely differentiated, their manufacture involves the use of a much greater variety of specialized intermediate goods (and intra-industry trade consists largely of trade in such intermediates)."* The delocalization of different production stages all over the world feeds up the intra-trade volume. *"So that it is not surprising to see Germans driving Hondas while Japanese drive BMWs"* (Krugman, 1995).

We note also that "Category 93" (Cat.93: other machinery apparatus and appliances, engines, parts thereof) is in the first position both in import and export. This category contains manufacturing products characterized by high economies of scale (e.g. engines production): the Krugman's statement finds evidence in this observation.

Now, let's focus on European imports from the eight country-partners.

CHINA:

- 45% of European Union import from China belongs to "93 category" (Cat.93: "Other Machinery Apparatus and Appliances, Engines, Parts Thereof"). Since China is a labour-intensive country, products within this category should be components.

INDIA:

- 26% of European Union import from India belongs to "96 category" (Cat.96: "Leather, Textile and Clothing"). India is labour-intensive and it's consequently specialized in labour-intensive goods like clothing.

JAPAN:

- 43% of European Union import from Japan belongs to "93 category" (Cat.93: "Other Machinery Apparatus and Appliances, Engines, Parts Thereof").

UNITED STATES:

- The first category is the "93 category" (Cat.93: "Other Machinery Apparatus and Appliances, Engines, Parts Thereof"), but the percentage is relatively smaller (29%).

AUSTRALIA:

- European imports from Australia are peculiar and they don't include the "common categories" seen in table 16.
- Coal (Cat.21) is the first imported product: Australia is the country with the highest number of coal mines in the world.
- "Category 93" (that usually covers high percentages) is only 7,52%.

21 - COAL	31,38%
45 - NON-FERROUS ORES AND WASTE	13,38%
93 - OTHER MACHINERY APPARATUS AND APPLIANCES, ENGINES…	7,52%
12 - BEVERAGES	5,74%
97 - OTHER MANUFACTURED ARTICLES	5,69%
99 - MISCELLANEOUS ARTICLES	5,47%
89 - OTHER CHEMICAL PRODUCTS	4,68%
56 - NON-FERROUS METALS	4,29%

Table 16 - European Imports from Australia

Source: Eurostat

BRAZIL:

- Brazilian exports don't reflect the usual product categories.
- The firsts Brazilian category is "Iron Ore" (Cat.41) with 17,3%; Brazil is one of the main iron ore's producers (with Australia and China).
- The second category, "Stimulant and Species" (Cat.13), covers 9% of total export.

CANADA:

- The first Canadian category is the common "Cat.93" with 17,4%.
- Canada is rich in raw materials. The country exports in Europe minerals, iron and other metals (Cat. 63 "Other Stone, Earths and Minerals" (8,8%); Cat. 41 "Iron Ore" (7,5%); Cat.45 "Non-Ferrous Ores and Waste" (6,8%)).

RUSSIA:

- Russia is rich in energy resources and minerals; most of Russian exports toward Europe include these available resources (see Table 17).

31 - CRUDE PETROLEUM	48,7%
32 - FUEL DERIVATIVES	14,2%
33 - GASEOUS HYDROCARBONS, LIQUID OR COMPRESSED	8,5%
56 - NON-FERROUS METALS	3,5%
99 - MISCELLANEOUS ARTICLES	2,4%
21 - COAL	2,2%
81 - BASIC CHEMICALS	2,2%
52 - SEMI-FINISHED ROLLED STEEL PRODUCTS	1,1%
63 - OTHER STONE, EARTHS AND MINERALS	0,9%
97 - OTHER MANUFACTURED ARTICLES	0,8%
51 - PIG IRON AND CRUDE STEEL; FERRO-ALLOYS	0,7%
93 - OTHER MACHINERY APPARATUS AND APPLIANCES…	0,6%

Table 17 - European Import from Russia

Source: Eurostat

- 48% of the European imports from Russia is “Crude Petroleum”.
- In this case, the common category 93 cover only 0,6% of the overall import.

4. Transportation Patterns Determinants

4.1 Model Purpose

We started our thesis describing the nowadays context of the globalization and the consequent fundamental role of trade costs for economic balance of companies and countries. We explained how globalization could have been caused by the reduction of transportation costs along the 20th century. However, costs didn't change a lot trade sensitiveness to distance so that many economists jumped to the conclusion that "distance is not dying". The literature "discusses" a lot about the actual meaning of trade sensitiveness to distance today. We spoke about the predominant role of time and all of the possible reasons why it looks so important to go towards "timely trade". For sure, one possible solution to be fast in deliveries is to trade with close countries; however, airplanes freight can overcome the time constraint by permitting the same responsiveness, while trading with the "ideal" far partner (we can assume that the best variety is owned with higher probability by far rather than neighbour countries). Even if costs are going down, plane usage still implies massive extra-cost, so that it's clear from the "preliminary investigation", trade by air represents only a minority of overall trade (that minority composed by the most valuable products). We want to further investigate on the role of fast mean of transport today. Our goal is to understand which are those key-features allowing air transportation, to assess impact of the identified determinants, to understand

better trade-off between extra-costs and speed advantage in the globalized context.

We will conclude trying to identify and to generalize findings concerning trade in general, giving our contribution about the “distance puzzle” in gravity model.

Let’s start interviewing “Artsana S.pA.” in order to get “real” evidence about international transportation patterns. This is the starting point for our hypotheses to be tested.

4.2 Artsana S.p.A. Interview

Artsana S.p.A. is an Italian leading company operating in baby apparel, toys, nursing, healthcare and cosmetic products. Its major brand names are:

Baby-Care → Chicco, Boppy, PRENATAL

Health & Beauty Care →Pic, SERENITY, LYCIA, CONTROL

Artsana is a multinational firm that operates and sells in many countries of the world. Thanks to its international attitude, it is a perfect "real example" for our studies: through the interview we aimed at obtaining a clearer view about international trade by understanding the actual decisional processes about transportation.

We interviewed a manager of the Logistic (import-export) function. We prepared in advance open questions to be asked. In the following paragraph we report all the submitted questions and their answers.

QUESTION 1: In which Extra-EU countries does Artsana ship (or receive) goods?

"Artsana is a multinational firm. Its clients are worldwide distributors. In particular the two brands "Pic" and "Chicco" are sold in many Extra-European countries."

QUESTION 2: Which are the preferential transportation modes?

"Toward Extra-EU countries, the most of goods (about 90%) is shipped "by sea". Airplanes are used just in urgency and in particular situations. Transport "by road" is largely used for East-Countries like Russia, Ukraine and Belarus. Since our customers are mainly big retailers, their orders fill one or more containers."

QUESTION 3: Which are the main determinants that you consider during the decisional process for the transport mode selection? For example: time, costs, distance, fragility…

"*Cost in general is the main determinant. Time becomes relevant just in particular situations in which the airplane becomes the only one solution*."

QUESTION 4: So, in which cases do you decide to select the "air shipping" mode?

"Airplanes are "exceptions". We adopt air shipping mode for <u>urgencies</u> (e.g. stockouts) or for the <u>fashion apparels and shoes</u> sector (for example the Chicco apparel goods). We use mainly "sea shipping" for the health-care sector.

QUESTION 5: Is there any constraint for fragility?

"No, fragility is not a relevant aspect. In the summertime, we used refrigerated container for some particular goods like plasters and thermometers in order to

assure a temperature of 10°C; these special containers can be shipped by sea and they are more costly than a traditional container."

QUESTION 6: From a technical point of view:

A) Which are the types of ships used? Bulky, Cointainer, Ro-Ro…

"*Ships are mainly containerized. Since we are a big company, we operates with the major steamship companies like MSC (Neapolitan company with headquarter in Geneva) and COSCO (a very competitive Chinese company). In general, these companies are associated to local shipping agents for assuring the multimodal transportation (Sea+Road)."*

B) Which types of containers are used for the sea shipping?

"We use all of the traditional types of container: 20 foot, 40 foot and 40' High Cube."

C) Which types of airplanes are used? Cargo planes or Passengers planes?

"It depends on the company. Our partners are usually large logistic companies using air-cargo shipping; but that's not the rule."

D) Which types of pallets are used for the air transport?

*"For air shipping, it's necessary to use particular pallets in plastic or in fumigated wood; the traditional wooden pallet cannot be used if you ship by air. We use the traditional pallet dimension 80 cm * 120 cm with a maximum height of 200 cm."*

QUESTION 7: Do you try to fill entire container (in order to get saturation) or do you ship also small quantities? In case of small quantities, are the logistic companies in charge of combining different cargoes? Or is it rather up to you?

"In general, we usually order full containers for import in order to optimize the shipping saturation. About export, even our customers order in general full containers."

QUESTION 8: According to cost accounting:

A) Which are the main cost determinants for shipping a container "by sea"? Harbor fixed costs, distance, insurance…?

"Surprisingly, the main voice of cost is many times the transport from the harbor to the firms. Hence, the total cost depends on the distance between the firm and the harbor. For example, in these days we are shipping a container from Como to Honk Kong and we pay about 1000 Euros for road transport but only 250 dollars for the sea shipping. So, nowadays the national road transport is much more costly than the sea shipping in case of China-Italy trade. This because China needs empty containers to export towards other countries; hence the export toward China becomes very convenient.

Let's consider the export of a 40 foot container from Como to Honk Kong. In general the sea shipping cost structure is defined in the following way:

***FOB (Free on Board) costs** are all those costs related to the shipping from the firm to the departure harbor. In general there is a cost per kilometer: the farer the harbor, the higher the costs. (800/1000 Euros)*

***THC (Terminal Handling Charge) costs** are all those costs associated to the loading/downloading of cargo into/from the ships. Each port has its own tariff that can depend on the harbor characteristics. (150/160 Euros)*

***CUSTOM DUTY**: it depends on the countries and on the heterogeneity of the cargo. In Italy, the custom duty is about 50 Euros for a mono-*

product container (with only 1/2 custom voices); the duty can be up to 200 Euros in case of more than 20 custom voices. (Max 200 Euros)

***OCEAN FREIGHT**: that's the cost for sea freight (harbor to harbor). Nowadays, the freight towards Honk Kong costs only 250. A "normal" reasonable price should be 800 US Dollar. (250 USDollars)*

The other costs related to THC and to road shipping are in charge of the customer from the destination harbor to the final destination (FOB costs).

In case of import, Artsana buys in FOB: transportation costs are in charge of Artsana starting from the national harbor (THC and FOB costs).

All of the costs and responsibilities policies follow the rules established by "INCOTERMS 2010".

B) So, it seems that the effective sea shipping freight is not very influential…

It depends on the moment and on the specific case. In this moment, we have an "anomalous" situation for the export towards China. But situation can be different from case to case. For example, the route Italy-Argentina has OCEAN FREIGHT costs equal to half of total shipping costs. In general real OCEAN FREIGHT is about half of the total costs involved in shipping.

C) Could you give some costs estimation for different destinations?

Starting from Italy we have about the following complete tariffs:

- *North and South America = 3000/3500 Euros*
- *Oceania = 4000 Euros*
- *North Africa = 500 Euros*

- *South Africa = 4000 Euros*

D) According to "air shipping", which are the main costs determinants?

Air shipping tariff are not fixed and they can change suddenly for many different reasons. Air transports is divided in two main types: "Direct" and "Consolidate". With a "Direct Flight" you assure a flight without any stops in intermediate destinations and delivery times are short (one or two days). In case of "Consolidate Flight", air companies book cargo spaces for a flight taking off at the end of the week. In general, the "Direct Flights" are much more expensive but faster.

QUESTION 9: let's compare times now: how long does it take for ships or planes to get China starting from Italy?

The effective "sea shipping time" is about four weeks. While the "air shipping time" is much shorter; the plane takes at latest 1 day to reach the final destination; however we must add 24 hours of "quarantine" for the anti-terrorism laws. There is a new widespread multimodal transportation that is called "sea-air". It's common that ships leave any Asiatic harbor directed to Dubai (it takes about eight days). From Dubai, cargoes reach the final destination by plane. This particular "solution" is a kind of "intermediate transportation mode" between the "pure" air and sea shipping (both in terms of time and costs).

QUESTION 10: According to the intra-EU trade. Which kinds of transport mode are used?

The most of the goods are transported "by road". Nowadays, it's possible to reach by truck all the main European cities in 2 or 3 days. Airplanes are very rarely used.

We transport goods towards Russia almost only "by trucks". We have recently opened a new plant in Russia and we transport goods "by road" in 8/10 days.

QUESTION 11: How has the transportation changed in the last decade? (The role of time, Customer's habits, Costs?)

Transportation has not substantially changed. Oil price's oscillations have affected the transportation tariffs over time; we must consider that the "fuel charge" tariff has continually changed.

The liberalization of "sea shipping" tariffs is one of the most relevant things changed in the last 10 years. Until four years ago, there were the so called "Conferences": meetings in which the main sea shipping companies agreed about tariffs level. Nowadays, there's tariff liberalization so that firms must face higher transportation costs uncertainty; moreover, many small sea freight companies couldn't sustain the stronger competition and they had to merge together.

Even from the customer point of view change has occurred. In the past, customers used to plan big quantities orders in advance. Nowadays, customers order only what they need (many orders and smaller quantities) due to the higher uncertainty and risk of the current context. Artsana faces this scenario by increasing its own stock level and by using plane for urgencies (we must be ready for any request). The "groupage" method is another solution facing this problem: when customers order small quantities, the logistic companies combine different small cargos to saturate containers; however, the "groupage" method is much more costly than a "full container" shipment and it can take much more time.

QUESTION 12: Finally, do you think that the choice about transportation mode is relevant from a strategic point of view?

The choice of the transport mean is strategic: we need accessible costs and proper delivery time.

4.3 Trade Costs "Drivers"

There are many factors affecting trade costs and determining how goods move. As already said, it's not only a matter of monetary **costs** of transport. For instance it's very important to understand which is the role of **time** and **risk** in trade.

According to the literature and to preliminary observation of data, we identified four categories of factors that influence concurrently trade and transportation patterns: **Product** characteristics, **Time** demand, **Distance** among trade partners and countries' **Connectivity.**

We will call them "**trade dimensions**" and the thesis' goal is to understand their integrated impact over transportation patterns.

We start now this chapter, analyzing these "trade dimensions" one at the time and we will describe their key-features (the concrete "**drivers**" that we will use for the model). Finally, we will formulate hypotheses to be tested empirically through a regression analysis.

4.3.1 Product

Trade involves very different kinds of goods. Any product has its own characteristics and transportation choices depend on them. For example, no one will ever transport a mineral like "coal" by air because coal has very low value/weight, low "fragility" and it doesn't lose any value over time. Looking at the "Eurostat dataset", 93% of coal was imported by sea, 6% by rail and less than 1% by air (data refers to EU27 imports in weight, year 2010).

There are three main drivers related to "product dimension": **value/weight**, **density** and **fragility**.

VALUE/WEIGHT

Value/weight is defined as the ratio between value (Euros) and weight (tons) of goods. This indicator can be considered a valid approximation for goods price, at a high level of data disaggregation.

According to the economic literature about this topic, **air freight is used for goods with a high value/weight ratio**. James Harrigan (2005) states that "*air shipment is much more likely for goods that have a high value/weight*"; David Hummels (2007) shows that "*the probability of selecting the air transport instead of the sea shipment steps up at the increase of the value/weight of good*".

From the preliminary analysis of data (Eurostat db), it's clear that goods moved by air have much higher value/weight than goods moved by sea (considering the average value/weight of goods imported by the EU for different means of transport, "air value/weight" is 107 times bigger than the sea one: Figure 20).

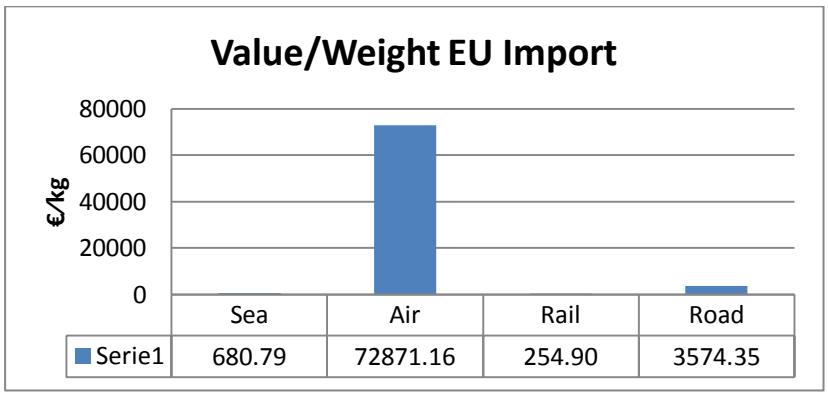


Figure 20 - Value/Weight of European Import by Mode of Transport

Source: EUROSTAT

There are three main reasons why value/weight has a positive correlation with the likelihood of using air freight.

The first one is the **"percentage impact" of transportation over FOB prices**: if goods have high value, transport costs have lower impact on final prices (if transport cost is 10 €/unit, it has different impact if companies have to

charge it over the price of a book or of an I-pad.). On the contrary, goods with very low value weight, as the coal we mentioned, are hardly transported by air because unit costs would be much higher.

The second one is about the **opportunity cost**: goods in transit represent a capital that is not investible. The opportunity cost is traditionally computed as (Harrigan, 2004):

*OC = [value of the good] * [daily interest rate (risk free rate)]*[days in transit]*

Opportunity cost increases for higher values of goods, and airplanes can reach almost any destination in a few days or less while it takes weeks for ships to freight containers. Therefore, the opportunity costs is much higher in case of sea shipping because the capital is chained in transit for longer time: it's very expensive (in terms of opportunity costs) to transport high-value goods by sea.

The last point is about **safety** of means of transport. Risk of pilferage is higher for sea shipping, and if high-value products are damaged, the loss is higher (it's less risky to ship valuables by air).

Air shipment is more likely for goods with higher value/weight ratio.

DENSITY (MASS AND VOLUME)

Density is defined as the weight of each unit of volume **[kg/m^3]**.

Years ago, transport rates were computed according to goods **weight**. Heavier goods were always charged more than lighter products. This policy resulted in low-dense goods to be not very profitable for logistics companies: transport was paid a little because of low weight of products, while occupying a lot of the freight space.

Volume capacity must be considered a significant constraint in shipment pricing. Cargoes space availability has a limit and carrier companies should concurrently fill all of the weight and volume availability to maximize profits. This ideal best

solution is called **Optimal Cargo Density** (by definition: "OCD is *the density of freight that would occupy the total available cubic capacity of a truck while simultaneously reaching its cargo mass limits*" - International Transport Forum, OECD department, "Moving Freight with Better Trucks", 2010).

Nowadays, companies use to calculate a "density lower boundary". If the density level is higher than the boundary, price is computed according to the weight. But behind this density level, there is a premium price that changes according to volume.

A significant number of transportation companies adopt "**Volumetric Weight Standards"** set by "The International Air Transport Association" (IATA) to define the boundary.

Here IATA standards for domestic and international air shipment:

International transport = 166.08 [in^3/lb] or 6000 [cm^3/kg]

Domestic transport= 193.76 [in^3/lb] or 7000 [cm^3/kg]

Many other companies (especially for ground delivery) define their own volumetric weight factors.

The logistic "giant" FedEx such as the German company DHL use to compute a "**Dimensional Weight"** to set fair rates. DHL reports procedures to be used on its official website[11]: *"You should calculate DIM on every shipment to its actual rate, and use the greater of the two to determine the shipping cost. Here the general formula → Length * Width * Height in centimeters / 6000 rounded up to the nearest 0.5 kg".*

Considering that air transport is the most expensive option, low-density freights, far from the optimal cargo density, involve much higher losses if delivered by air.

[11] DHL official website: http://wap.dhl.com/serv/volweight.html

Moreover, airplanes have very strong limitations in terms of space. It could be very difficult to "exploit economies of scale" while shipping by air "volumetrically big" goods (such as cars). It's impossible to load a single airplane with a high quantity of very large goods, so that most of times, it will be necessary to ship twice or three times to get all the units that are needed in the period unit. This results in fixed costs duplication and increased "cost per unit".

Air shipment should be more likely for relatively small and not very light goods.

FRAGILITY & SECURITY NEED

Fragile freights need particular care during transport because of the high risk of damage during jolting and handling. Airplanes are usually safer than ships because of tighter controls and more structured handling procedures. Moreover, air shipment has much shorter transit time: the likelihood of damages is lower because the "risk exposition" is shorter ("Shipping and Incoterms", Practice Guide, UNDP).

Air shipment should be more likely for very fragile goods.

4.3.2 Time

There are many reasons why customers are willing to pay higher price to get what they need quickly.

From a mere financial perspective, no one should even consider the idea of shipping goods by air, because this option is invariably more expensive than any other way of transporting goods (Evans and Harrigan, 2003).

However, things don't work that way. Since the plane is the most expensive, but also the fastest mean of transport, planes are widely used for international trade: in 2010, EU27 import by air was about 20% of the total moved value (Eurostat db). **This demonstrates the great importance of time in business decisions.**

David Hummels (2001) states that one of the most valid reasons why time is so important in business is the "depreciation of goods": *"Depreciation captures any reason that a newly produced good might be preferable to an older good"*. This phenomenon can be the result of intrinsic spoilage of products (like the perishability of fresh products).

The time dimension is even more important if there is the need of extreme **punctuality** to **synchronize** the different phases of the production process or if end-users require very **high levels of service**. Since consumers are more and more demanding, and products have shorter life-cycle, it's clear that the "trade timeliness" (Harrigan, 2001) becomes an important source of competitive advantage for companies.

PRODUCT LIFE CYCLE

The global competition, the technological innovations and the **transition from the mass market to the full customization** had an impact on demand and life-cycle of products.

Nowadays, consumers change tastes very often and companies must be able to follow faster production rhythms. It's much more difficult to forecast demand today, since products are complex, offer range very wide and demand extremely volatile (Hummels, 2001).

Transport time is a "critical success factor", especially for those products that have short life-cycle. Hummels (2001), states that "*with long lapse of time between production ordering and final sales, firms may face a mismatch*

between what consumers want and what the firms have available to sell". Obsolescence of goods is one of the main risks linked with delivery: if it takes too much for a product to reach the market, it might lose its value to the customer.

The clothing industry can be very "time-sensitive". Fashion is "in evolution" and there are products whose lifetime doesn't exceed four weeks. If it takes three weeks to deliver something that is sellable for 4 weeks total, it can be "troubles": sea shipping isn't a good solution.

That's the case of "Zara". Its success is directly connected to the concept of "fast fashion" based on the "QR strategy" (Quick Response). The main goal is to maintain a quick inventory turnover. They replenish goods in stores twice in a week and no product is exposed in shops for longer than four weeks. It's necessary to minimize shipping time to implement this strategy

To conclude**, products with short lifecycle can't be transported by sea**. It's better to choose the faster air shipment despite its greater cost because shipping time can't take the most of the product life.

In the last decades average life-cycles became shorter and shorter. If this trend won't change, trade timeliness will become a more and more important "success factor".

Air shipping should be more used for goods with short "life-cycle".

PERISHABILITY RISK

Some goods can't sustain lengthy shipping times because of their nature itself. Dairy products, cut flowers, newspapers require rapid deliveries, otherwise their value decays totally and they become pure waste: no one is interested in reading daily news if they refer to days ago and fresh milk expires after a few days.

Perishable products need also specific condition during transport such as a particular temperature and humidity. Most of times, perishable goods are also

"fragile" so that a lot of attention is required during carrying. Air shipping is very likely to be the most convenient solution for this kind of products.

According to a research by "Merge Global Consulting" (carried out for Lufthansa), "perishables" constitute around 15% of the total air cargo. This data is expected to grow 7,1% annually. A possible explanation is that air freight is more accessible today, than ever before. Countries can choose to import perishables from far away producers (by air) because the premium price for fast delivery is counterbalanced by lower "FOB prices".

SERVICE LEVEL AND NEED FOR TIMELINESS

In many industries, companies must guarantee a given level of service by **contract**. It means that penalties occur if they delay deliveries. In some sectors, very brief lead-times are forced by law, or Governments "sponsor" fast deliveries.

Let's think about the pharmaceutical industry. If a company is developing a new vaccine, (for instance against the new seasonal influence) the urgency of the situation requires prompt time of reaction. Once the company discovers the new vaccine, they can't wait weeks to get the components and start production. Air shipping is the only solution to satisfy the required service level.

That's exactly what happened in 2009. A small chemical Australian company discovered "Relenza": a quite effective vaccine against the H1N1 "swine flu". After having sold the license to the pharma giant "GlaxoSmithKline", the "Tamiflu" was produced in huge quantities, requiring a massive usage of air freight cargo.

Nowadays, high service levels are source of **competitive advantage** for any company in any industry. Quick and punctual deliveries result in higher sales and few wastes: if customers can choose between many similar alternatives, they are more willing to purchase something from the "fastest producer".

The need for "timeliness in trade" is increasing worldwide as a consequence of two main concurrent trends: the ever-growing **volatility of demand**, and the **vertical specialization** of the supply chain.

High volatility of demand means that it's very tough to understand in advance what and when to produce. If companies can't rely on accurate forecast, they need very **high "SS"** (safety stock) to avoid stock-out.

After Toyotism and its new managerial paradigm, companies started adopting "the just in time" approach. According to "JIT", companies should produce what they need only when they actually need it because any **extra-stock (SS)** is source of waste. The production process and the delivery must be fastened, with smaller production lots and more frequent replenishments. This results in strong need for timeliness**.** Coordination and punctuality become fundamental since it's not possible to "pull the process" if the components required don't reach the production plant on time. **On the one hand, JIT solves the "volatility problem" by eliminating waste in the production process; on the other hand, it stresses the need for timeliness in production and transportation.**

The world stock of "FDI" (Foreign Direct Investments) rose 12% yearly on average from 1980 to 2010 ("UNCTAD-stat data", Figure 21). Vertical specialization among different countries, means that components or WIPs must be transported abroad to meet the following stage of the supply chain. The different "actors" involved in the process need to synchronize activities because customers' expectations about service level are the same with or without "internationalization". **On the one hand FDI permits to exploit labor comparative advantages; on the other hand it stresses the "need for timeliness" in transportation activities.**

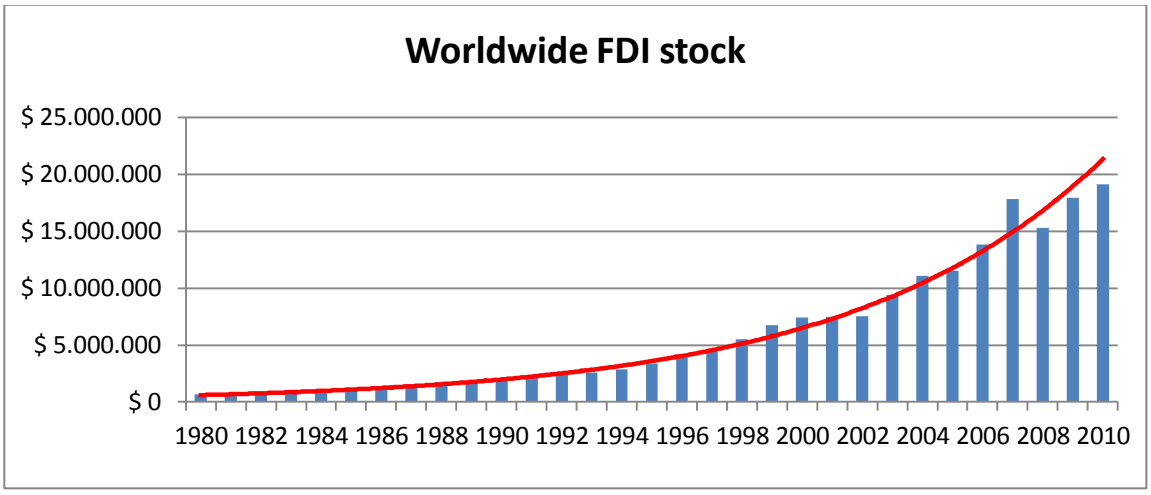


Figure 21 - Worldwide FDI Stock

Source: UNCTAD

Being air shipping faster and faster, it represents a solution to undermine the tradeoff between the need for timeliness in industrial demand and the "internationalization" of the supply chain (Hummels, 2007).

Harrigan and Venables (2004) translate "need for timeliness" in "need for proximity" stating that risk of delay, and uncertainty in delivery, should push towards a clustering of supply chain, rather than a mere cost driven internationalization of firms.

Summarizing, the increase of vertical FDIs and the more and more uncertain demand make speed and timeliness two critical success factors; expansion of the JIT technique and international supply chain clustering are possible improving solution. Even though, air shipping is often the more effective way to implement these strategies and reach the market goals.

Air shipping is more used if customers claim higher levels of service.

4.3.3 Distance

There is a strong link between distance and trade costs since the most considerable part of cost items (freight fees and fuel consumption) increase linearly over distance. Transportation rates are sometimes assessed according only to distance to be covered. *"Because of the measurement of transport costs, the geographic distance has dominantly be used as a proxy for the transport costs"* (Kiyoyasu Tanaka, 2009).

For reason of simplicity, some shipping companies use a "step-wise" approach to establish transport fares (figure 22).

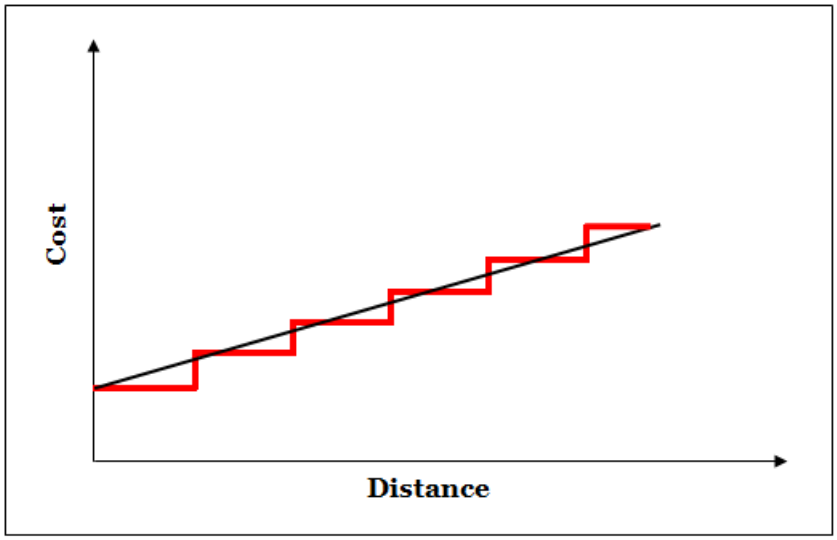


Figure 22 - "Step-wise"Transport Fare

Source: The Geography of Transport Systems

DIRECT COSTS OF DISTANCE

In economics, they use to underline the difference between fixed and variable costs. "*Transport service incurs a number of costs such as labor, fuel, maintenance, terminal, roadway, administrative and others. These costs can be divided into two categories. Firstly those which vary with services or*

volume, i.e. variable costs, and secondly those which do not, i.e. fixed costs. Fixed costs in transportation refers to costs like roadway acquisition and maintenance, transport equipment and administration costs. Variable costs, on the other hand, refers to costs fuel and labor and equipment maintenance". (Kapoor and Kansal, 2005).

In traditional accounting, they consider "variable" those costs that change with production volumes: the larger the quantity produced, the higher the variable costs. In trade, costs change according to haulages' volume, weight and distance to be covered. **In this section, we will consider variable costs in relation to distance**.

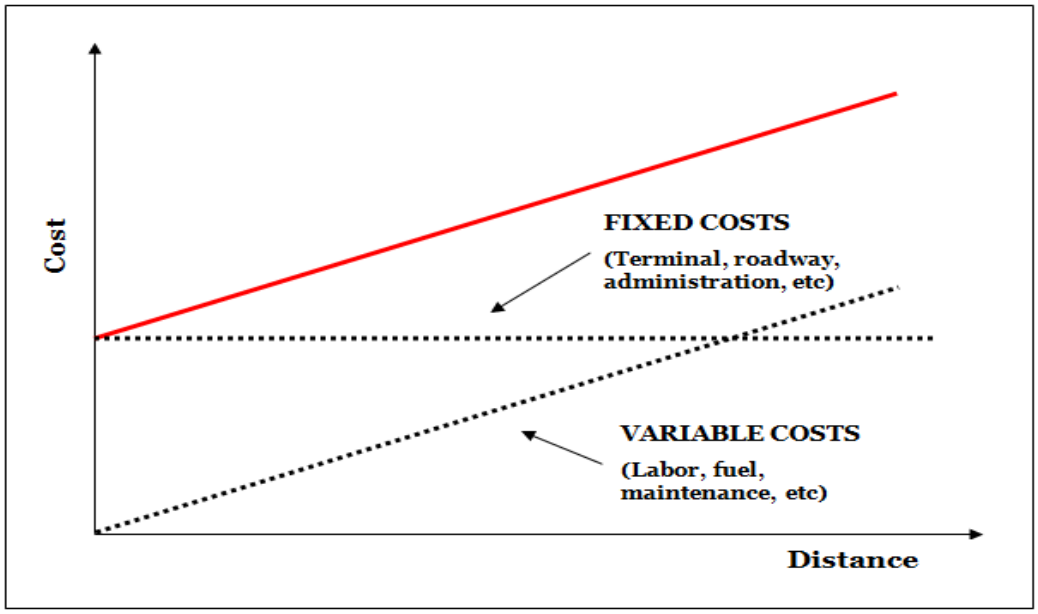


Figure 23 - Transportation Cost Structure

Source: http://www.valocchi.eu/logistica/trasporti.htm

Different transport systems have different cost allocation. Sea and air shipping present very high fixed costs since they require expensive "ground operations" like cargo loading and unloading, warehousing and terminal fees (see Table 18).

TRANSPORT MODE	FIXED COSTS	VARIABLE COSTS
Rail	High	Low
Road	Low	High
Sea	Very High	Very Low
Air	Very High	Very High

Table 18 - Fixed and Variable Costs by Mode of Transport

Source: Valocchi

A research by "CONFETRA" (Confederazione Generale Italiana dei Trasporti e della Logistica), states that the incidence of "ground operations" can reach 40% of total for companies shipping containers (Table 19).

COST	INCIDENCE
Overheads	15 %
Equipment (containers) expenses	15%
Ship expenses	30%
Ground handling expenses	20%
Harbor fees and taxes	20%

Table 19 - Cost Structure of "Sea Shipping"

Source: Confederazione Generale Italiana dei Trasporti e della Logistica

"Variable costs" differ a lot between air and sea transport: airplanes must sustain the highest variable costs while "the ships' expenditure" grows very slowly over distance.

Let's try to represent the different costs functions in a graph (Figure 24).

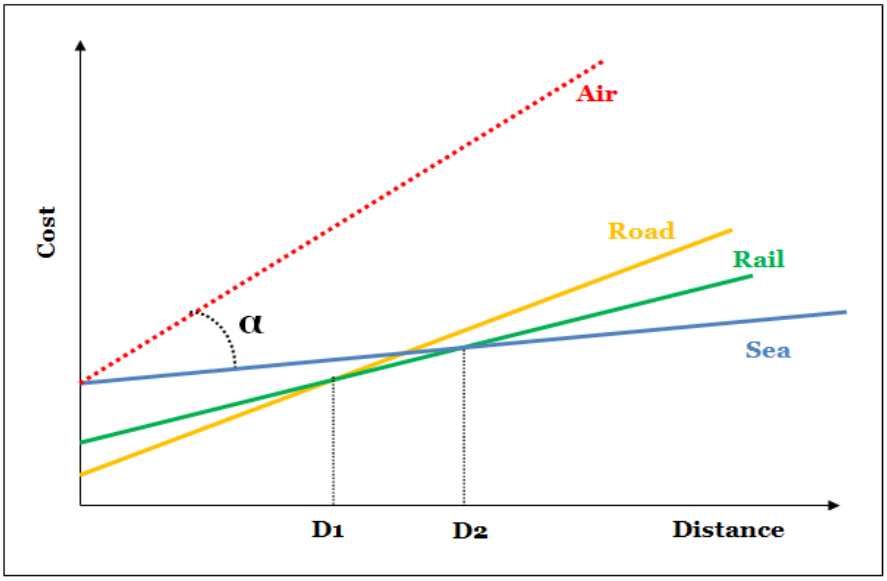


Figure 24 - Cost Functions by Mode of Transport

Source: The Geography of Transport Systems

Road, rail and sea functions are taken from "The Geography of Transport Systems" website, while we draw the "air function" according to the data available on the Internet. Planes' fixed costs are supposed to be similar to those of ships, while variable costs are supposed to be much higher. The slope of the line ("α coefficient") could be different because we didn't test it empirically.

We can do some easy considerations looking at the graph. Planes are never convenient from a "direct cost" perspective, while ships become the cheapest mean after "D2" (the long distance): **from a financial point of view, sea shipping should be the only mean used for transoceanic routes.**

Ships and planes seem to be equally disadvantageous for "short distances" since they have very high fixed costs. In the reality, planes' variable costs (drawn as linear to distance) are different somehow. Fuel consumption is much higher during take-off and landing so that a part of the "fuel expenditure" is actually an extra fixed cost for planes (there is a taking-off and a landing in any route). **This makes air shipping even more expensive for "short distances".**

DISTANCE AND TIME

Airplanes' advantages don't rely on direct costs, but on other performances such as time or safety. Actually trucks don't seem to be slower than planes for distances behind 1000 km: the lower "net transit time" of planes is counterbalanced by much longer "fixed times" (it takes much longer to go through maintenance and controls activities).

Here (table 20) we report expected time for different distances by sea and by air.

Route	Distance	"Air time"	"Sea time"	"Road Time"
Genoa - Catania	925,0	1,0 day	2,0 days	<2 day [12]
Genoa – Tokyo	6193,0	2,0 days	12,0 days	/
Genoa – Perth	13685,0	3,0 days	24,0 days	/

Table 20 - Time Difference

Source: TNT and sea.rate.com simulator[13]

[12] For the Genoa - Catania route, you might use trucks. In ideal conditions, it can take less than 14 hours (transit-time) to reach Catania by car (130 km/h on the freeway). Nevertheless, you have to consider fixed times (e.g. for loading/unloading activities), and you have to consider that big trucks can't go over 90 km/h on the freeway. By law, a driver can drive 10 hours straight (with a 1-hour break in the middle); then he has to rest for 11 hours before starting again. If there are two drivers, it can take less than one day to deliver by truck. Otherwise, it takes almost 2 days.

[13] Sources: http://www.searates.com/
http://www.tnt.it/it/itsolutions/tempi_consegna.shtml

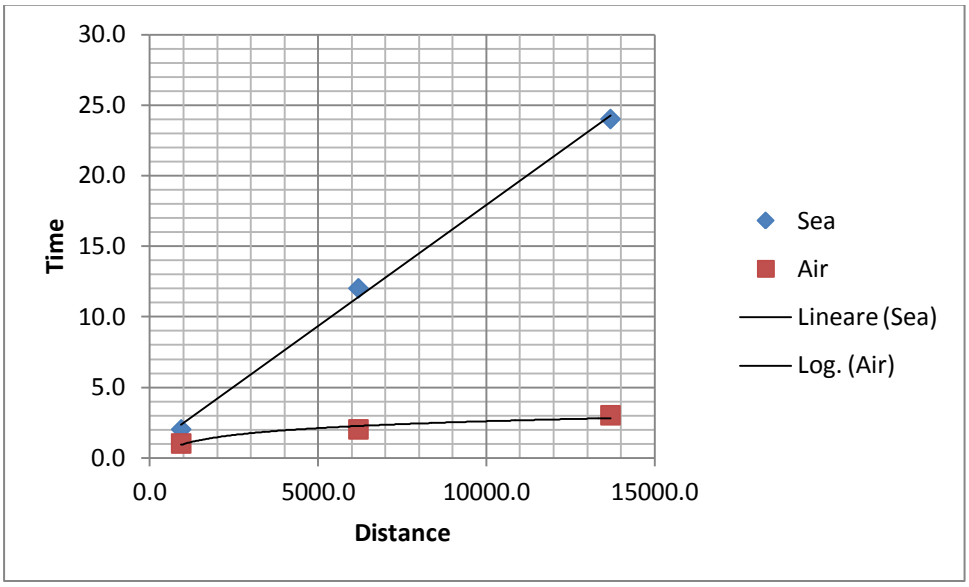


Figure 25 - "Sea" and "Air" Shipping Time

Source: TNT and sea.rate.com simulator

The functions drawn are based on data in table 20. Such a few data can't have statistical evidence, but it's enough to make some considerations.

- "air function" is not linear since "time slacks" are less than proportional to distance: planes have higher "fixed times".

- "sea function" is much steeper: **planes are faster and faster for increasing distances**.

Since there could be a relevant kilometric difference between "air route" and "sea route", the great differential in delivery times between the two transport means is caused not only by a speed difference, but also by a sensible difference in kilometric hauls. On the one hand it takes 2 days (considering even fixed times) to deliver in Tokyo and 3 days to deliver in Perth by air. On the other hand, it takes 29 days to go in Tokyo and only 24 days to go in Perth by ship. **This to underline that "nautical distances" can change transportation patterns.**

Ships can reach destinations on the coast only. It means that "inter-modal transport" is required if the destination is in the hinterland. This makes delivery even slower and sea shipping less convenient. We will discuss this topic in the next chapter.

Airplanes are rarely used for short routes (<1000 km).

Trade "uses" almost exclusively planes and ships for long distances (>7000 km).

"Air share" keeps increasing with distance, over the "long distance boundary" (>7000).

4.3.4 Country Connectivity

The capability of a country to transport goods abroad is one of the essential elements to survive in the current global market; an isolated country with inadequate infrastructures, outdated technologies and inefficient administrative systems can't trade easily with the rest of the world.

The level of "connectivity" has a great influence over the choice of the transport mode; for example, if a country doesn't have reliable or accessible airports, the air shipping will result very difficult.

In general, there are three main drivers affecting the connectivity level of a country: **the available infrastructures, the level of technology and the required bureaucracy**.

AVAILABLE INFRASTRUCTURES AND TECHNOLOGICAL LEVEL

The availability of infrastructures is an important structural constraint for trade. The location of airports and maritime ports, their accessibility and their "connectivity capacity" are related to transportation patterns.

Countries governments are in charge of decisions about investments and projects in transport infrastructures, and their means are directly connected to the economic level of the country.

Since developing countries can invest less than developed ones in infrastructures, poorer countries have usually more "accessibility problems".

Let's think about Turkey and Italy for a numerical comparison. According to the "OECD database" (about investments in sea ports infrastructures), although the available coastlines are quite the same for these two countries (7458 km for Italy and more than 8000 km for Turkey), the investment in maritime ports is widely higher in Italy (Figure 26).

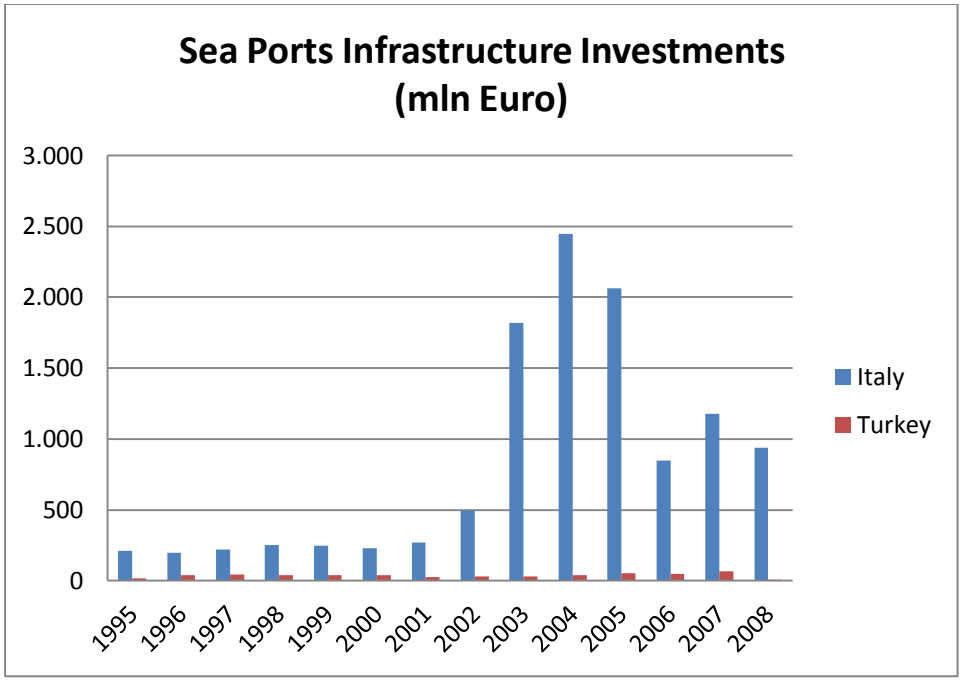


Figure 26 - Italian and Turkish Investments in Port Infrastructures

Source: OECD

The difference between maritime and aviation infrastructures development is what most influences the choice of the mode of transport. If a country has few and underdeveloped airports, but a good endowment of harbours, air shipping becomes less likely.

It's not just a matter of density: even the infrastructure technological level has a great relevance. Regarding maritime ports, there are substantial performance differences between a conventional break-bulk terminal and a containerized one; containerized harbours allow to better exploit the volume capacity and it requires much less handling time "per ton". On the other hand, containerized harbours require more advanced technologies and managerial tasks that are not available in all countries (see Table 21).

Conventional Port	**Containerized Port**
Small terminal surface	Large terminal surface
Direct transhipment possible	Indirect transhipment
Limited mechanization and automation	Advanced mechanization and automation
Improvisation in terminal operations	Organization and planning

Table 21 - Coventional and Containerized Ports

Source: The Geography of Transport Systems

Being time advantage the main driver that leads to air shipment, technologies available at air terminals must guarantee quick loading and unloading operations; in order to assure quick freighting operations, the airports must be provided with all the cargo facilities like the X-ray scanning and adequate warehouses.

All over the world, there are few airports dedicated only to freight traffic. They are hubs for logistic companies like Federal Express (Memphis airport) and UPS (Louisville airport); these central hubs are mainly located in the most congested

traffic points of North America, Pacific Asian zone and Western Europe (Figure 27).

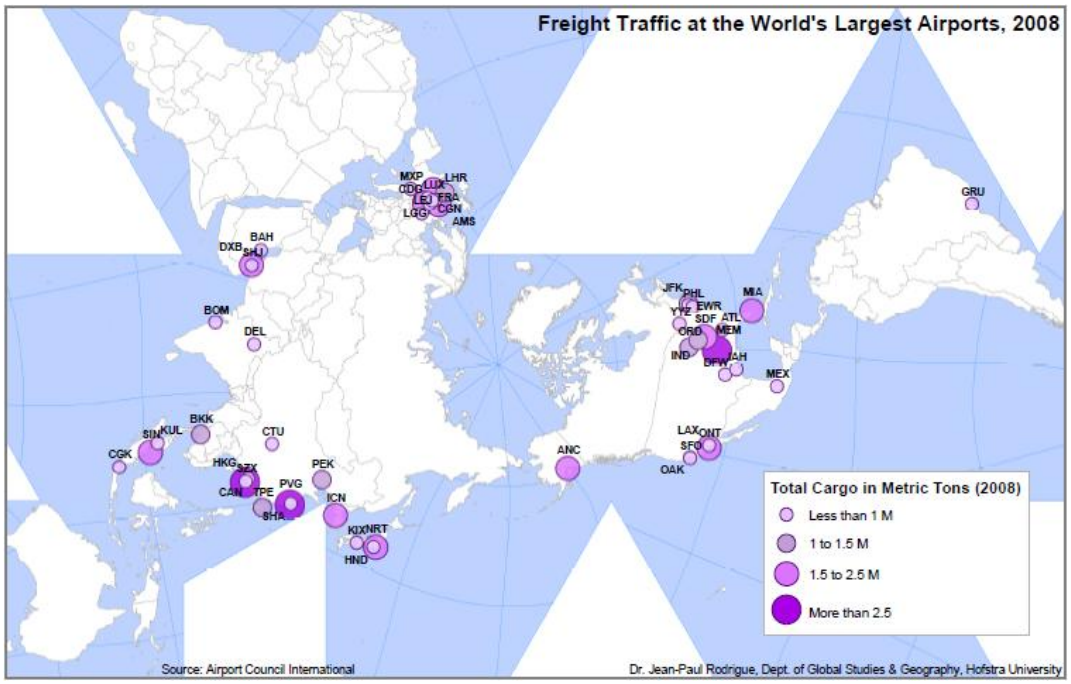


Figure 27 - Freight Traffic at the World's Largest Airports, 2008

Source: Airport Council International (2008)

BUREAUCRACY

Regarding bureaucracy, there are no substantial differences between air and sea shipping.

The Bill of Landing (B\L) is the main document for the sea freighting. It contains all the relevant information about the cargo (ship code, markings, types of goods, number of packages, etc.); in the same way, the Air Waybill (AWB) is the principal document for the air shipping. Other documents such as bank documents, Custom Clearance documents, port and terminal handling documents and cargo airlines transport documents can be required according to specific country regulations, and to characteristics of cargoes (dangerous goods, perishable goods, etc).

4.4 Data Gathering, Statistical Adjustments and Dataset Creation

Every empirical analysis needs reliable data to get reliable results. Our aim is to create a consistent and complete database that includes representative data for each variable.

Sometimes it's easy to get economic data. There are many international institutions that collect information about the main economic measures for each country (e.g. GDP, demographic information, international trade, etc…); World Bank, OECD, Eurostat, IMF and UNcomtrade are just few examples of such international institutions. These institutions are labeled as "reliable"; however, there could be few errors and imperfections that could distort the statistical results. For this reason, we take particular care to the "outliers' management".

Other times, data are not easily available and it's necessary to use proxies or "qualitative solutions". In this case, the risk of data meaningless and distortion is higher; hence, it's important to check the reliability of data sources and to adopt a systematic approach to assign qualitative measures.

In our model some variables are directly measurable like GDP, distance, infrastructures' information and value/weight. On the contrary, other variables like time sensitiveness, product density and transportation costs are not readily available in the traditional data sources.

In the following paragraph we explain how we have tackled data gathering.

4.4.1 Point of Reference

First of all, we clarify the international boundaries that we selected for our empirical analysis.

According to the preliminary data analysis (c.f. 3.6 “Focus on European Trade”) analysis’ reference point is the international trade (import and export) of the European Union (27 countries). We choose the following 11 countries as country partners:

- Australia
- Brazil
- Canada
- China
- India
- Japan
- Morocco
- Russia
- Switzerland
- Turkey
- United States

The selecting criteria for countries partner are the same that we used in the preliminary analysis:

- Country’s geographical position: we want to cover different distances from the EU27. In this way, the dataset is uniform according to the distance variable and the observations are independent and identical distributed (i.i.d.).
- Country’s economic relevance: we select relevant countries that are very “open” to international trade.

We consider the NSTR product classification proposed by Eurostat. In order to be more precise we decided to take the 3-digit level with all the 175 product categories.

In total we get 3400 observations. In Appendix D we report an extract of the database.

4.4.2 Dependent Variable

Our dependent variable "Y" represents the probability to use the air shipment to trade a given product with a given country.

There are several datasets available created by international institutions that collect data about international trade like Eurostat, Unctad, Uncomtrade, WTO, World Bank, etc...; each dataset has its peculiarities in terms of completeness, level of detail, time series and reliability. The selection of the source of data must be first of all functional to the current analysis; under this perspective we select the Eurostat database[14]. In the section "External Trade" there are reliable and complete data related to bilateral trade, divided by mode of transport and reported in terms of value and weight (tons).

The assumption is that amount of bilateral trade shipped by air during 2010, from EU to a given partner (or vice-versa), represents the probability to select the "air mode" for that product, for that country.

$$Y\% = \Pr(air|product\ i; country j) = \frac{Value\ of\ trade\ by\ air_{i;j}}{Value\ of\ total\ trade_{i,j}}$$

[14] From http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/; "External Trade" → External trade detailed data → EXTRA EU27 Trade Since 1999 By Mode of Transport (NSTR) (DS_022469)).

Since Switzerland and Russia share a boundary with the European Union and they have a direct "paved road" connection, we decided to develop a corrected dependent variable called "Yspeed%". This variable (Yspeed%) indicates the probability of selecting the faster mode of transport and it can be useful for several analyses.

The basic concept is that airplanes are the fastest mode of transport, but for Russia and Switzerland even trucks "look competitive" since average transportation times by road are comparable to those we have by air (look at the time difference evaluation below).

$$Yspeed\% = \mathrm{Y}\% \qquad \forall\, j \neq \textit{Switzerland or Russia}$$

$$Yspeed\% = \Pr(air + road|i;j) = \frac{Trade\ Value\ by\ air_{i;j} + Trade\ Value\ by\ road_{i;j}}{Value\ of\ total\ trade_{i,j}} \qquad \textit{when j=Switzerland or Russia}$$

4.4.3 Independent Variables

Value Weight

Value Weight is a product characteristic. It can be seen as a proxy for the price (cost, insurance and freight price) of goods. For any observation considered (for any Y), we just computed the ratio between the overall value and the overall weight traded.

For instance Value/Weight for product i, country j (in import or export) is:

$$VW\ [€/Tons] = \frac{[Tot.Value\ (product\ i, country\ j, import/export)]}{[Tot.Weight\ (product\ i, country\ j, import/export)]}$$

Distance

Distance is an important country characteristic. It's widely used in the economic literature as descriptor of bilateral trade, and many times transportation cost is seen as a proxy of distance.

It's difficult to define a single given distance representing the real "remoteness" for two countries to trade. Trade can involve different sites within the country, and these sites can be far enough to distort analysis.

We found well-adjusted data on the "CEPII website": *"The CEPII has built and made available two datasets providing useful data for empirical economic research including geographical elements and variables. A common use of these files is the estimation by trade economists of gravity equations describing bilateral patterns of trade flows"*[15].

CEPII computes two different types of distance. The simple one is calculated with the "great formula circle". Since the Earth is a sphere, the shortest distance between two points is represented by an arc. It is actually the simple distance between capital cities or economic centers. The weighted distance is an average of more distances between all of the relevant sites within the two countries. The weights are defined according to the population density.

We select this second alternative because it can well represent the "real remoteness" between two countries.

The relative distances are calculated considering Germany (Berlin) as the centre of the European Union and they are expressed in kilometers.

After the first analysis we decided it was necessary to go through some adjustments concerning distance for Switzerland and Russia. The Cepii weighted distance from Germany can't be representative for these two countries.

[15] Cepii website: http://www.cepii.fr/anglaisgraph/bdd/distances.htm

Switzerland is actually in Europe and the relatively different distance towards different countries in the EU27 can affect the actual "real remoteness" parameter. A similar reasoning can be done concerning Russia. It's not always the distance to Germany to be relevant for trade (Russia has a contiguous boundary with Europe).

To solve the problem we computed a "second level weighted average", exploiting trade and distances (for Russia and Switzerland) towards the 27 countries in the EU. We computed an average distance weighted on relative bilateral trade (values in Euros).

We report in Appendix D the spreadsheet relative to Switzerland. Switzerland remoteness to Europe is relatively higher than weighted distance to Germany (659 km instead of 543 km). Similar results for Russia: from 2655 km to 2703 km.

Connectivity

Connectivity is a country variable. It should indicate how much it's "easy" to trade "by air" or "by sea" with other countries. We defined two different indicators. The first one refers to the airport availability. We measured it through the ratio between number of major commercial airports[16] (we summed all of the principal airports for the countries in the analysis according to international flow) and country surface. Since the final number was very small, we multiplied it by 10^6.

The second indicator refers to sea and road connectivity. Usually harbor capacity is very large and it doesn't represent a constraint. But under the assumption that it's impossible to get the final destination by sea without using intermodal

[16] Source of data: http://www.flightstats.com/go/Airport/airportsOfTheWorld.do

transportation (while we assume air shipping to be "door to door" because distance to be covered by road is negligible if compared to total distance), the percentage of paved road[17] within countries becomes an important connectivity indicator for both of road and sea transportation mode. A low percentage of paved roads is much more disadvantageous for sea shipping than for air shipping since the distance to be covered by road is very different in the two cases.

Gross Domestic Product

GDP is a country measure. It is used to estimate the economic size of a country. By definition GDP is the value of final goods and services produced by an economic system (e.g. a country) in a given period (e.g. one year).

We considered GDP as a relevant control variable: since GDP can be directly related to the other independent variables (e.g. GDP is positive correlated to the country's connectivity) the regression's results can be distorted.

In order to compare the GDP of different countries we included GDP in "Purchase Power Parity" conditions (PPP) expressing the value in Euros. The data source is the World Bank database[18].

However, GDP is not a good indicator for the country's wealth. Population dimension and income distribution must be considered. For example, Japan's GDP is similar to the India's GDP, but the level of wealth of the two countries is very different. For this reason, we decided to add the GNI/capita to the regression analysis. The GNI (Gross National Income) "*is the sum of value*

[17] Sources of data: http://data.worldbank.org/indicator/IS.ROD.PAVE.ZS http://www.nationmaster.com/graph/tra_roa_pav_of_tot_roa-transportation-roads-paved-of-total

[18] Data Source: http://siteresources.worldbank.org/DATASTATISTICS/Resources/GDP_PPP.pdf

added by all resident producers plus any product taxes (less subsidies) not included in the valuation of output plus net receipts of primary income (compensation of employees and property income) from abroad"[19]. The GNI/capita spread the overall country's income to the population and it can better indicate the wealth of the country.

Time Sensitiveness

"Time Sensitiveness" is a product peculiarity. It tries to measure the impact of time on the different product categories.

In the traditional economic literature, time dimension is a very recent variable and there are few empirical models that include time.

According to a Deardorff's conceptualization (2001), time must be introduced in the economic models as "depreciation": "*…goods depreciate over time, becoming worth less the longer the time that elapses after their production*".

The basic concept is that some products depreciate faster than others; hence, some products are more time sensitive than others. Our aim is to detect the "depreciation speed" of different product categories.

Since direct measures about time sensitiveness are not available, we decided to adopt a qualitative approach. First of all, we defined three main factors influencing the time sensitiveness: Product Life Cycle, Perishability and "Need for Timeliness". Product life cycle expresses the obsolescence's risk; perishability indicates the physical deterioration (or the complete value loss) of the different product categories; "need for timeliness" measures the importance of speed and timeliness along the supply chain.

[19] Data Source: http://data.worldbank.org/indicator/NY.GNP.PCAP.PP.CD

For each of them and for all of the 168 product categories, we assigned a score from 1 to 3 considering the following scales that we modeled (Table 22):

PRODUCT LIFE CYCLE	
1	Long life cycle (no obsolescence risk)
2	Medium life cycle
3	Short life cycle (high obsolescence risk)
PERISHABILITY	
1	No perishable (long lasting product)
2	Medium perishability
3	Fast perishability
TIMELINESS	
1	Low need for timeliness
2	Medium need for timeliness
3	High need for timeliness

Table 22 - Qualitative Scale for Time Sensitiveness

When the categories were too heterogeneous or when the information about the product characteristics were not available, we omitted the score (in this way, we overcame the risk of having data distortion).

In the end, in order to get a single variable for "time sensitiveness" we made a weighted average of the three parameters; we assign 0.4 to "perishability" and 0.3 to "product life cycle and" and "timeliness": in this way we gave more relevance to "perishability". The assumption is that when a product deteriorates, it becomes pure waste, while when a product becomes obsolete (or when there is a components' delivery in delay) there is just a partial loss of value because it's possible to sell the product at a lower price.

In the Appendix C we report the complete table with the scores for each category.

Density

Density measures the weight of each unit of volume. Despite its relevance in international trade, product density has been ignored in the most of the economic literature.

Data about volumes are not available and the main international institutions record international trade just in terms of value and weight.

Also in this case, we faced this obstacle by introducing a qualitative analysis. Our idea is to use the "Specific Weight" as a good proxy of product density.

The specific weight (also known as "unit weight") is available for several materials[20]; nevertheless, the main problem we tackled is the high heterogeneity of some product categories: for such categories it's impossible to estimate an average density and we omitted the density value. For example, it's easy to assign the unit weight to the "Raw Sugar" (Cat. 111 of NSTR) because the product is very homogeneous; on the contrary, it's very difficult to assign a value to the "Electrical Machinery, apparatus, appliance and engines; parts thereof" (Cat. 931 of NSTR): this class contains very different products including both small and light microchips and bulky and heavy refrigerators.

In the Appendix C we report the complete table with the values for each category.

[20] Specific Weights: http://www.asiinstr.com/technical/Material_Bulk_Density_Chart_T.htm

From Distance to Time Difference and Delta Fares

In statistics, one of the main causes of distortion is the variable's omission. Omitted-Variable bias (OVB) occurs when a regression model neglects one or more important causal factors. The distortion is created when the model "compensates" the missing variables by over- (or under-) estimating one of the other factors.

That's the case of distance. Distance includes concurrently the effects of two omitted variables: the time required to reach the countries and transport fares. Since this omission could cause a distortion in the distance coefficient itself, we decided to add two further independent variables to the regression model.

The dependent variable expresses the probability of choosing air shipment rather than sea freighting. For this reason, we didn't simply report the air shipping time and the air fares, but we computed the time difference and the delta fares between "air" and "sea" transport modes.

In the regression model we didn't include together all the three variables, but we substituted the distance with the other two factors. The reason is that the inclusion of all the three variables in the same regression could cause a distortion of "**imperfect co-linearity**". There is imperfect co-linearity when two or more variables are highly correlated, that is when a linear function is able to explain the correlation between variables (there is instead "perfect co-linearity" when one of the independent variables is a linear combinations of the others). Imperfect co-linearity (unlike the perfect co-linearity) doesn't trigger great problems to the OLS estimators, but it can bias at least the coefficient of one of the variable.

Here beyond, we expounds the data gathering and manipulation of both the variables.

TIME DIFFERENCE

This variable indicates the average difference between "air transportation time" and "sea transportation time" for each country-partner. The difference is expressed in terms of days.

We collected "air shipping" data from the TNT web-simulator[21]: the web application allows us to estimate delivery times from/to any destination. The system estimates delivery times according to the first available solution. So, if there is no availability for the next flight (the freight is already booked), the system addresses automatically your order to the first free shipping cargo. This resulting in time distortion (we just need the "net time" for transportation). To solve this problem, we made queries distant in time (delivery to be done in one month).

For "sea shipment" we queried the Sea.Rates's website[22]: the shipping time is estimated for a 14 knots fast vessel.

We used Rotterdam as the final European destination for our evaluations, because Berlin (the reference site for the distance computation) is land-locked, while Rotterdam is the hub-harbour for the most of European trade.

The "time difference" variable looks adept at describing time gaps for all of the countries except for Switzerland, Russia and Turkey. Switzerland is land-locked and it's quite unlikely that sea shipping is widely used for trading with the EU. Moreover, "road shipping" is much faster than the sea one for the most of the European destination (from Zurich, Moscow or Istanbul). So that, for these countries we will consider the trade off between transportation time "by air" and "by road". To estimate transportation time "by road" we simply used the "Google

[21] Data Source:
http://www.tnt.com/wwtt/wwttinput.do?navigation=1&respLang=en&respCountry=gb

[22] Data Source: http://www.searates.com/

Maps" application[23] (or ViaMichelin application[24] for Russia); we corrected data by a coefficient representing mandatory breaks for drivers: we referred to Italian regulation that imposes to stop 1,5 hours after 4,5 hours driving, then they can drive other 4,5 hours before stopping for 11 hours. Finally we added a coefficient representing fixed times.

Switzerland (time for a transportation from Zurich to Rotterdam)

- Net time = 8 h

  Transportation time = 8 h + 1,5 h = 9,5 h

  Delivery time = 9,5 h + 2 h = 11,5 h

Russia (time for a transportation from Moscow to Rotterdam)

- Net time = 27 h

  Transportation time = 4,5 h + 1,5 h + 4,5 h + 11 h + 4,5 h + 1,5 h + 4,5 h + 11 h + 4,5 h + 1,5 h + 4,5 h = 53,5 h

  Delivery time = 55,5 h + 2 h = 55,5 h

Turkey (time for a transportation from Istanbul to Rotterdam)

- Net time = 30 h

  Transportation time = 4,5 h + 1,5 h + 4,5 h + 11 h + 4,5 h + 1,5 h + 4,5 h + 11 h + 4,5 h + 1,5 h + 4,5 h + 11 h + 3 h= 67,5 h

  Delivery time = 67,5 h + 2 h = 69,5 h

[23] Google Maps application: http://maps.google.it/

[24] Via Michelin application: http://www.viamichelin.it/

COSTS DIFFERENCE

Costs Difference measures the ΔFares between air and sea shipping.

Costs difference = "Air Fare" – "Sea Fare"

To capture the effect of distance, it's not enough to consider the time gap for the different routes, but we have also to consider the costs' perspective.

First of all, we obtained data simply querying web-applications of two logistics companies: the idea is to compute "Δfares" just changing the routes and fixing other parameters such as size and weight that are asked by software (Table 23).

Route	**"Sea Fare"**	**"Air Fare"**	**"Δ Fares"**
Destination 1	…		
Destination 2	…		

Table 23 – "Fare Difference (1)" Definition

We adopted the same approach we used in the "Technical Chapter" (C.f. 3.5 Transport Fares: Empirical Evidence) considering FedEx Corporation[25] to estimate air shipping fares and IQ Global Logistic Corporation[26] for sea freighting. Since the IQ Global Logistic Corporation operates just to/from the USA and the sea fares don't vary a lot for different distances, we adopted an average sea fare considered constant for any destinations (209 €/pallet). We labeled the new variable "Fares Difference 1".

According to an interview to a manager of Italmondo S.p.A.[27] (Italian leading company in international transportation), we had the confirmation that:

[25] FedEx web-application:
https://www.fedex.com/ratefinder/home?source=gh&cc=it&language=it

[26] IQ Global: http://www.iqglobal.com/

[27] Italmondo: http://www.italmondo.com/it/index.html

"*…Sea fares don't depend on distance but our rate tables are divided by destination zones and costs are sensible to imposed tariff and harbor's fixed costs*".

On the one hand, this first indicator is very simple to be computed; on the other hand, we are considering standard arbitrary measures of freight (pallet of 50 kg, 50cm x 50 cm x 50 cm) for queries on the Internet that could be source of mistakes: the "costs proportions" could change a lot according to the real weights and volumes to be traded: delivering a freight 10 times bigger than the standard measures we considered, it's unlikely that costs will be perfectly ten times higher (due to the presence of economies scale).

Moreover, as already said airplanes and ships have obviously weight and volume constraints, this resulting in costs that vary not only with distance or routes, but even with product density. There is actually no direct link between the probability of using airplanes and the product density itself. On the contrary, density changes directly transportation costs, this resulting in different relative convenience for different transport systems. While considering separately the variables "density" and "costs difference", there could be coefficients distortion.

We built a second indicator "Fares Difference 2" trying to combine the two variables effects: for any route to be queried, we introduced different parameters according to different products densities (Table 24).

Route	**Density**	**"Sea Fare"**	**"Air Fare"**	**"Δ Fares"**
Destination 1	Density 1	…		
	Density 2			
	Density 3			
	…			
Destination 2	Density..	…		

Table 24 - "Fare Difference (2)" Definition

This second indicator is more structured, but there is still a strong limit: we keep taking into consideration a fixed freight volume for any delivery equal to 50 dm^3 (while changing the relative weights according to density). However, the first constraint we talked about isn't solved.

For reasons of simplicity, we rounded the actual densities into 9 categories. We report in table (Table 25) the weights used, corresponding to 50 dm^3:

DENSITY CATEGORY [kg/dm^3]	VOLUME [dm^3]	WEIGHT [kg]
0.5	50	25
1	50	50
1.5	50	75
2	50	100
3	50	150
…	…	…

Table 25 - Weight of 50 dm^3 depending on Density

In the Appendix C we report the Specific Weights of each product category.

To improve robustness, we elaborated a third cost difference indicator: "Fares Difference 3". It's more likely that bigger "unit of transport" are used for both of air and sea cargo transportation: that's the reason why we decided to take into consideration the "size of trade" for any observation.

The probability to use airplanes (the dependent variable) is built as the percentage of import/export by air compared to overall trade for that product towards that country. Hence, the cost difference must be weighted over the overall amount of trade. Using the previously described indicator, it seems that the cost difference is the same delivering 50 tons or 50 kg of potatoes.

The idea behind this approach is to compute how much it would cost to transport the amount of import/export summarized in the dataset observations entirely by sea or by air.

The basic assumption is that 20/40 foot standard containers are always used for sea shipping, while the standard 120*160*200 cm pallet is commonly used for air shipping.

Using the density coefficient already described above, we can translate the amount of import/export for any observation even in terms of volume.

Firstly, we evaluated "sea fare" while computing how many standard transport units (40 foot container) would be necessary to transport entirely by sea the amount of products summarized in the observations. Hence, we simply solved the following inequality:

(Weight of trade/weight unit transport) > (vol. of trade/vol. unit transport)

The number of container to be delivered corresponds to the higher ratio. For instance if 10 containers are needed to transport the weight of products in the observation and 12 container are needed to transport the volume of products in the observation, then we take 12 containers, reflecting the more restrictive constraint.

Once we have the number of 40foot standard container to be used to transport everything by sea, we can compute the cost for sea transportation while only estimating cost to deliver one standard transport unit at different destinations (and then multiplying for the number of containers needed).

On the Internet, it is very difficult to find these data, so that we had to ask for information to different logistic companies to understand costs. We got interviews with the logistic giant UPS, with the "local" group Artsana (that has to deliver in the most of the countries we consider in the analysis) and with "New Transport": the logistic company in charge of managing distribution for Artsana. So, we could get information about both of net costs and fares.

Artsana gave us prices concerning sea shipping for the different destinations; prices that take into account all the main transportation cost voices including average insurance.

About air shipping, we went through a similar process while comparing actual weight for observations with their volumetric weight. Of course, the bigger value between weight and volumetric weight reflects the more restrictive constraint between weight and volume.

We didn't compute the number of standard pallets needed by singles observations since we used the UPS tables to evaluate "air fare". However, these tables describe fares applied for standard pallets transportations. So, it was enough to multiply weight of the "entire observation" (or volumetric weight depending on the higher value found) by the right fare applied by UPS. Fares in table vary according to the weight of goods; tariffs are divided in 8 different destination zones and each zone has different costs.

Once we got cost estimations for sea and air fares, it was enough to compute the slack between them.

The last "adjustment" we made reflects the **saturation level**. It is quite unlikely to fill completely containers and pallets with any type of product (any size, feature and shape), especially under the assumption that everything is transported using the same type of containers and pallets (our basic assumption).

We assessed "the saturation level" for any category fixing a coefficient ranging from 0 to 1 (we report the saturation level for each category in the Appendix C). This adjustment increase the level of arbitrariness on the analysis, but it's somehow necessary to make the indicator "reliable". Let's think about the "live animals" and "rice" categories. It's quite unlikely to have a same given level of saturation freighting live animals or rice!

So, we have:

Fare3= "Air Fare" – "Sea Fare"

Sea Fare = (#Containers)(Cost/Container|Destination)*

#container = max [(Trade Weight)/(Weight Unit Transport); (Trade volume)/(Volume Unit Transport)]

*Air Fare =Weight_Adj * (Fare|Destination Zone)*

Weight_Adj = max [Weight of Trade; Volumetric Weight of Trade]

*Volume = (Weight) /(Density*Saturation Level)*

Similarly to what we did for the time indicator, we had to correct measures referring to Russia, Switzerland and Turkey. Instead of using cost per container by sea, we considered cost per container by road because road transportation is easier and cheaper for these "neighbor" countries.

4.4.4 Outliers management

When dealing with regression analysis, the better the data gathering activities, the more consistent the results. For linear regressions not to be distort the observations must be i.i.d (Identically Independent Distributed), meaning that they are result of random sampling. Moreover, the presence of outliers should be very unlikely. From a statistical point of view, the sample distribution must have finite and not null curtosys, otherwise the tails of the distribution are so massive that there is high probability that some misleading and erroneous value is distorting statistical evidence.

From an operative point of view we tried to get rid of those observations whose values seem to be result of mistakes in gathering.

In particular, we implemented two main statistical corrections. First of all, we adjusted the dependent variable "Y%" by taking off all the meaningless observation that could bias results (even Eurostat could make errors in data gathering). We renamed the new "adjusted" variable as "Y%outliers".

Finally, we had to adjust the "Value/Weight" measures by developing Confidence Boundaries. There are some observation having import/export expressed in euros, but no import/export in tons since even only a few Euros of trade are reported, while weight under 0,5 tons are reported by Eurostat with "0". Obviously, we can't evaluate ratio having null denominator. Hence, we estimated an average weight of 0,25 tons, and we checked for likelihood of our estimation by considering the Student distribution of values.

We report in the Appendix D the detailed description of all the statistical corrections we made on data.

5. Empirical Analysis: Regression Modeling

5.1 Correlation Analysis

Once we computed all the operations that are required to obtain the final dataset in Excel (Microsoft Office application), we are ready for the statistical analysis. We chose the software STATA[28], an application fully dedicated to regression analysis.

The first thing we can do by means of STATA is to go through the "correlation analysis". It means that the software compares two variables at the time, to understand if there is any linear relationship among them.

The correlation coefficient is defined as follows:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_x \sigma_Y}$$

It represents the covariance of two variables. Its values range from -1 and 1, where a ρ equal to -1 means that the two variables are perfectly and negatively correlated, while a ρ = 1 means that there is perfect positive correlation: if ρ = 1, a positive variation of X implies a positive variation of Y.

If ρ = 0, it means that the variables are not correlated: if X assumes a positive value, it doesn't imply that Y will assume a positive value with higher probability.

[28] STATA: http://www.stata.com/

In other terms, the expected value of Y, conditioned by X, is equal to the expected value of Y "as is" → E (Y|X=i) = μY.

What's important to notice is that if the coefficient is low, it doesn't mean that there's no relationship between variables. Correlation measures just the level of linearity among variables.

In other terms, if there is linear relationship we will have a not-null ρ, but a not-null ρ doesn't imply the absence of any association between variables. We will discuss different types of relationships in the following paragraphs.

First of all, we calculated the correlation's coefficient between all the independent variables and the dependent variable Y (Table 26):

	CORRELATION	
REGRESSORS	Y %	Y Speed %
Value Weight	0.119	0.1144
Distance	0.2265	
Airport density	-0.0562	
Paved road	-0.0754	
GDP	0.0749	
GNI/pro capita	0.0664	
Density	0.1147	
Time Sensitiveness	0.2484	
Time difference	0.1988	
Fare difference 1	0.0942	
Fare difference 2	0.1758	

Table 26 - Correlation between Y% and all the Independent Variables

All the signs are consistent with our preliminary hypotheses except for "Airport Density" and "Fare Difference 1".

Looking at the correlation graph between “Y” and “Airport Density” (Figure 28), it’s clear that there is a distortion caused by the last “column” of observation (Swiss observations).

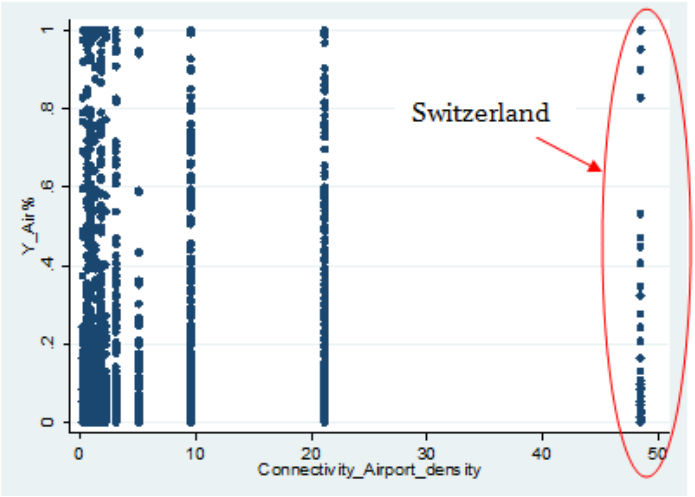


Figure 28 - Correlation between "Air Shipping" probability and Airports Density

Although Switzerland has the highest “Airport Density” ($48*10^{-6}$ airports/Km^2), Y is very low on average. The point is that Switzerland “can rely” even on roads to deliver fast. The same concept is valid for Russia. Hence, we made a correction by excluding Switzerland and Russia from analysis. Results confirm our hypotheses: the new correlation value is positive and equal to (0.101). We report the new graph in Appendix F.

Concerning “Fares Difference 1” coefficient, it’s seems surprising to get a positive result (although is not very high: 0.09): the probability of choosing airplane raises at the increase of “fare difference”. However, “Fares Difference 1” represents those parts of costs that vary with distance without concerning volumes and densities. Being distance coefficient extremely high and positive, being “Fare Difference 1” on average higher for longer distances, it results in a positive correlation.

We obtained interesting outcomes also for the value/weight variable. We made a first correlation analysis between “Y%” and “Value/Weight” and we got 0.119.

However, we can make two considerations looking at the graph (Figure 29). First of all, there are some “outliers”- values with very high Value/Weight but low “Y%”.

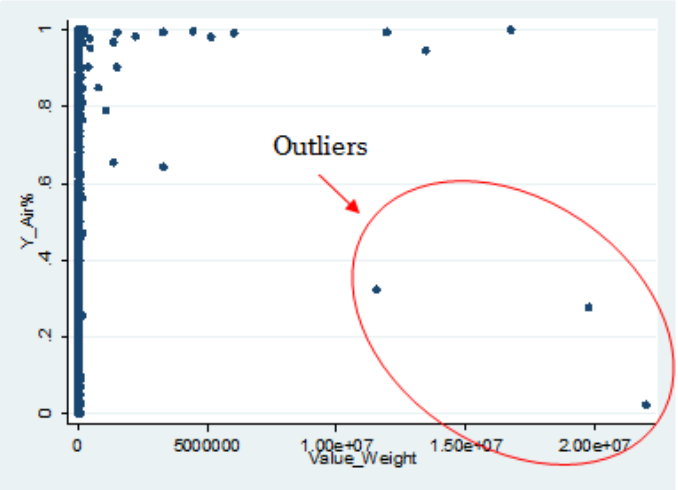


Figure 29 - Correlation between "Air Shipping" probability and Value/Weight

We overcome these problems by substituting “Y%” with “Yspeed%” (Figure 30): the distortion was caused by Switzerland and Russia that “use” also road for valuable goods. However, the correlation coefficient has not dramatically changed: from 0.119 to 0.1144.

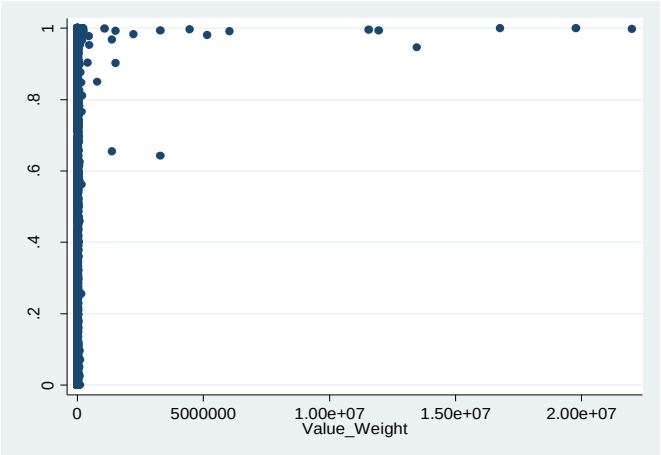


Figure 30 - Correaltion between "Yspeed%" and Value/Weight

Second, the most of the observations are concentrated in low-V/W range so that it’s difficult to understand Y trend on chart; in order to get a more visible graph

and to correct the coefficient from possible outliers, we decided to consider a shorter range of value-weights. Taking off V/W values over 10000 Euros, the relationship between the variables becomes more visible (Figure 31) and the correlation coefficient notably changes: from 0.1144 to 0.3.

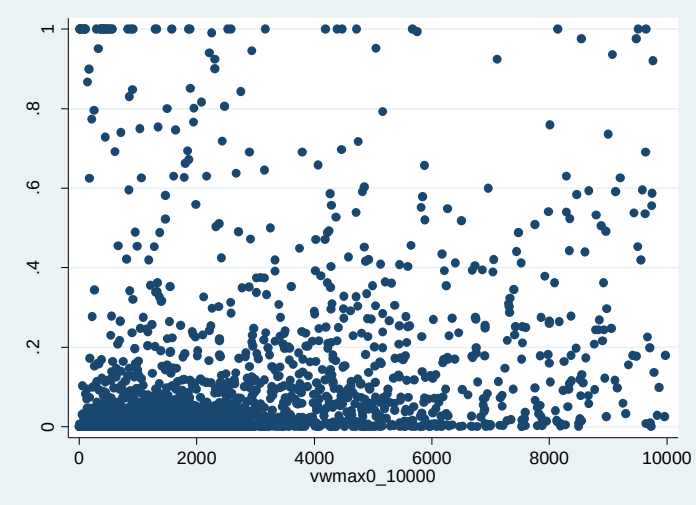


Figure 31- Correlation between "Air Shipping" probability and Value/Weight

In Appendix F we report all of the other correlation charts.

Once we completed analysis about variables correlation with Y, we tested correlation between independent variables. This analysis helps us to understand if there could be distortion in the regression's coefficients caused by the presence of "perfect" or "imperfect" **co-linearity**. "Perfect co-linearity" is not allowed by the OLS regression model (Ordinary Least Squares); while "Imperfect Co-linearity" isn't perceived by software, but it could cause distortion. Co-linearity definition has already been discussed above (c.f. 4.4.3 Independent Variables, "From Distance to Time Difference and Delta Fares").

We report results in the following table (Table 27):

REGRESSORS	CORRELATION
VW – Distance	-0.0322
GDP – Airport Density	-0.1020
GNI/capita – Airport Density	0.5793
GNI/capita – Paved Road	0.4136
Airport Density – Paved Road	0.6693
Distance – Fares Difference	0.5044
Distance – Time Difference	0.8106

Table 27 - Correlation among Independent Variables

All the coefficients are consistent with our hypotheses except for GDP and Air Density. We expected a positive correlation: the higher the GDP the higher the infrastructure level; however, GDP itself is not a very good indicator of country's wealth; testing the same correlation using GNI/capita, we got a more coherent result.

Concerning the co-linearity's problem, some other correlation coefficients are quite high. "GNI/capita" is correlated with both "Airport Density" and "Paved Road"; however, this "imperfect co-linearity" should not cause big problems in the regression (even if the coefficients could be concurrently distorted).

On the contrary, data confirm our hypothesis about the presence of co-linearity between "Distance"&"Time Difference" and "Distance"&"Fares Difference" (4.4.3 Independent Variables, "From Distance to Time Difference and Delta Fares"). Both of the coefficients are high, in particular the "Time Difference" one: including the three variables together in regression would cause distortion. We can consider the "Distance" variable by itself (that includes the causal effects of both time and cost), or rather to substitute it with the explicit "Time Difference" and "Fare Difference".

5.2 Regression Analysis

We start now the regression analysis. The goal is to develop a **regression model** both **representative** (the model must match reality) and **significant** (the model must be robust).

In order to get the optimal model, we adopt a **"step by step" approach**: starting from the easiest regression model, we improve it over time by introducing more complex variable's functions and interdependencies.

We assess model considering the **adjusted-R^2** indicator, the **signs' coherence** and the **significance level** of the coefficients' signs.

The **adjusted-R^2** [29] measures the capability of the regression model to fit the reality: it indicates if the variables are able to explain the available data. Unlike R^2, the adjusted-R^2 doesn't better when a new variable is introduced in the model, but it increases if and only if there is an effective improvement.

The **signs' coherence** is the consistency of signs (positive or negative) for any single variable.

The **significance analysis** is based on the following "hypotheses test":

H_0: the variable's coefficient β_j is null (variable doesn't explain data).

H_1: the variable's coefficient β_j is not null.

For each coefficient the system calculates automatically the "t statistics" and the statistical significance can be tested according to different confidence levels like 95% or 90%. For instance, a t-value equal to 1,96 means that there is a 95% of probability to refuse correctly H_0.

---

[29] Adjusted $R^2=1-\frac{n-1}{n-k-1}\frac{SSR}{TSS}$. $SRR=\sum_{i=1}^{n}\hat{u}_i^2$. $TSS=\sum_{i=1}^{n}(Y_i-\bar{Y})$.

5.2.1 STEP (1): Multiple Linear Regressions

The first step is the implementation of multiple linear regressions. This is the simplest regression model and we expected to get a rough data representation. The regression's coefficients are estimated by STATA according to the OLS method ("Ordinary Least Squares"). Variables' coefficient indicates the variation of Y (the dependent variable) according to a variable's unitary change:

$$Y = \beta_0 + \beta_1 X_1 \rightarrow \beta_1 = \frac{\Delta Y}{\Delta X_1}$$

We get the final multiple linear regression by inserting gradually all of the independent variables. We consider firstly "Y%" as the dependent variable. Then we try to correct the regression with "Y%outliers" and "Y%Speed". In the following tables (Tables 28, 29, 30, 31) we summarize the regression results considering "Y%" as dependent variable:

DEPENDENT VARIABLE "Y%"							
REGRESSOR		(1)	(2)	(3)	(4)	(5)	(6)
Value/Weight	Coef	$4.41*10^{-8}$	$4.41*10^{-8}$	$4.38*10^{-8}$			
	SE	$5.77*10^{-9}$	$5.77*10^{-9}$	$5.78*10^{-9}$			
	t	7.63	7.63	7.59			
Value/Weight Outliers	Coef.				$4.42*10^{-8}$	$4.42*10^{-8}$	$3.62*10^{-8}$
	SE				$5.18*10^{-9}$	$5.18*10^{-9}$	$5.44*10^{-9}$
	t				8.53	8.52	6.65
Distance	Coef.	0.000141	0.000141	0.0000144	0.0000125	0.0000105	$7.74*10^{-6}$
	SE	$1.09*10^{-6}$	$1.09*10^{-6}$	$1.3*10^{-6}$	$1.07*10^{-6}$	$1.68*10^{-6}$	$1.98*10^{-6}$
	t	12.98	12.98	11.02	11.61	6.23	3.91
Airport Density	Coef.	0.0005965	0.0004189	0.0000597	0.0003433	0.0003517	0.0001781
	SE	0.0003171	0.0004068	0.0004696	0.0003676	0.0003425	0.0004029
	t	1.88	1.03	0.13	0.93	1.03	0.44
Paved Road	Coef.		0.0170276	0.244537	0.0267931		
	SE		0.0244389	0.0237785	0.222274		
	t		0.7	1.03	1.21		
GDP	Coef.	$1.8*10^{-9}$	$1.58*10^{-9}$		$2.39*10^{-9}$	$2.36*10^{-9}$	$1.71*10^{-9}$
	SE	$9.59*10^{-10}$	$1.01*10^{-9}$		$9.17*10^{-10}$	$9.10*10^{-9}$	$1.07*10^{-10}$
	t	1.88	1.56		2.6	2.59	1.6
GNI/capita	Coef.			$3.13*10^{-7}$			
	SE			$3.55*10^{-7}$			
	t			0.88			
Delta Fares	Coef.					-0.0000384	-0.0000151
	SE					0.0000342	0.0000404
	t					-1.12	-0.37
Time Difference	Coef.						0.0014808
	SE						0.0007501
	t						1.97
R-Squared Adjusted		0.0681	0.0679	0.0675	0.0696	0.0696	0.0762
Root MSE		0.23996	0.23997	0.24003	0.21521	0.21521	0.21184

Table 28 - Linear Regressions (a)

DEPENDENT VARIABLE "Y%"						
REGRESSOR		(7)	(8)	(9)	(10)	(11)
Value/Weight	Coef	$4.4*10^{-8}$	$2.07*10^{-8}$	$2.06*10^{-8}$		
	SE	$5.81*10^{-9}$	$6.56*10^{-9}$	$6.53*10^{-9}$		
	t	7.56	3.16	3.16		
Value/Weight Outliers	Coef.				$2.1*10^{-8}$	$2.35*10^{-8}$
	SE				$5.63*10^{-9}$	$5.71*10^{-9}$
	t				3.73	4.11
Airport Density	Coef.	-0.0005218	0.0005946	-0.0014491	-0.0011623	-0.0011663
	SE	0.0003548	0.0004625	0.0004824	0.000421	0.0004275
	t	-1.47	-1.29	-3	-2.76	-2.73
GDP	Coef.	$1.63*10^{-9}$	$8.57*10^{-10}$			
	SE	$1.01*10^{-9}$	$1.31*10^{-9}$			
	t	1.61	0.65			
GNI/capita	Coef.			$2.4*10^{-6}$	$1.89*10^{-6}$	$1.93*10^{-6}$
	SE			$5.17*10^{-7}$	$4.53*10^{-7}$	$4.6*10^{-7}$
	t			4.63	4.18	4.21
Delta Fares	Coef.	-0.0000276	-0.0000155	0.0001533	0.0000939	0.0000946
	SE	0.0000376	0.0000493	0.0000607	0.0000534	0.0000542
	t	-0.73	-0.31	2.52	1.76	1.74
Time Difference	Coef.	0.0044455	0.0046639	0.0035468	0.00324	0.0032871
	SE	0.0004584	0.0005957	0.0006242	0.0005471	0.0005555
	t	9.7	7.8	5.68	5.92	5.92
Time Sensitiveness	Coef.		0.1504983	0.1502411	0.1611018	
	SE		0.0152769	0.0152008	0.0133499	
	t		9.85	9.88	12.07	
Time Sens-Dummy	Coef.					0.0900779
	SE					0.0100409
	t					8.97
Density	Coef.		0.138388	0.0139268	0.0137989	0.0122323
	SE		0.0020583	0.0020481	0.0017902	0.0018065
	t		6.72	6.8	7.71	6.67
R-Squared Adjusted		0.0004584	0.1117	0.1205	0.1476	0.1213
Root MSE		0.24155	0.24839	0.24715	0.21254	0.2158

Table 29 - Linear Regressions (b)

DEPENDENT VARIABLE "Y%"				
REGRESSOR		(12)	(13)	(14)
Value/Weight Outliers	Coef.	$3.85*10^{-8}$	$2.34*10^{-8}$	$3.83*10^{-8}$
	SE	$5.07*10^{-9}$	$5.72*10^{-9}$	$5.08*10^{-9}$
	t	7.59	4.09	7.54
Airport Density	Coef.	-0.0014546		
	SE	0.0003511		
	t	-4.14		
Paved Road	Coef.		-0.0287167	-0.0406185
	SE		0.0235766	0.0193679
	t		-1.22	-2.1
GNI/capita	Coef.	$2.44*10^{-6}$	$1.49*10^{-6}$	$1.89*10^{-6}$
	SE	$3.78*10^{-7}$	$4.3*10^{-7}$	$3.54*10^{-7}$
	t	6.47	3.46	5.33
Delta Fares	Coef.	0.0001181	0.0000864	0.0001021
	SE	0.0000444	0.000059	0.0000483
	t	2.66	1.46	2.11
Time Difference	Coef.	0.0030313	0.0034223	0.0032158
	SE	0.0004563	0.0005578	0.0004585
	t	6.64	6.14	7.01
Time Sens-Dummy	Coef.	0.1025584	0.0901037	0.1026264
	SE	0.0081083	0.0100556	0.0081265
	t	12.65	8.96	12.63
Density	Coef.		0.012211	
	SE		0.0018092	
	t		6.75	
R-Squared Adjusted		0.12222	0.1183	0.1117
Root MSE		0.20985	0.21032	0.24839

Table 30 - Linea Regressions (c)

DEPENDENT VARIABLE "Y%"				
REGRESSOR		(15)	(16)	(17)
Value/Weight Outliers	Coef.	$4.92*10^{-8}$	$3.83*10^{-8}$	$2.97*10^{-8}$
	SE	$6.12*10^{-9}$	$4.95*10^{-9}$	$6.89*10^{-9}$
	t	8.03	7.73	4.31
Airport Density ESATTO (no svizzera)	Coef.	0.0014589		0.0020153
	SE	0.0008037		0.0011752
	t	1.82		1.71
Airport Density	Coef.		-0.00141	
	SE		0.0003439	
	t		-4.1	
Paved Road	Coef.			-0.0179865
	SE			0.0306349
	t			-0.59
GNI/capita	Coef.	$2.11*10^{-6}$	$2.36*10^{-6}$	$1.49*10^{-6}$
	SE	$3.89*10^{-7}$	$3.71*10^{-7}$	$4.97*10^{-7}$
	t	5.43	6.36	3
Delta Fares	Coef.	0.0000968	0.0001087	0.0000531
	SE	0.0000451	0.0000436	0.0000618
	t	2.15	2.49	0.86
Time Difference	Coef.	0.0023467		
	SE	0.0004905		
	t	4.78		
Time Sensitiveness	Coef.	0.190659	0.1748597	0.1762095
	SE	0.0111555	0.0102786	0.0145183
	t	17.09	17.01	12.14
Density	Coef.			0.0143383
	SE			0.0019251
	t			7.45
TIME 0	Coef		0.0030189	0.0025846
	SE		0.0004336	0.0005881
	t		6.69	4.39
R-Squared Adjusted		0.1635	0.1593	0.1306
Root MSE		0.21191	0.20536	0.2292

Table 31 - Linear Regressions (d)

In order to get a more legible table, we differently colour the level of significance: "green" when the coefficient is significant at 95% (|t|>1.96), "yellow" at 90% (|t|>1.64), "red" when the coefficient is significant with lower confidence (we cannot refuse H_0).

In the first seven regressions we try to combine different variables considering all the independent variables that are related to the countries characteristics (distance, infrastructures, GDP, Δfares, time difference) and the "value/weight". The obtained results are poor: all the R^2-Adjusted are very small, and many variable's coefficients are not significant.

With regressions (8) and (9), we get very small improvements by adding "Density"; however R^2-Adjusted is small, some coefficients are not significant (regression (8)) and some signs are not coherent (e.g. "Airport Density" in (9) is negative).

Regression (10) is equal to (9), but we substitute "Value/Weight" with the adjusted "Value/Weight-Outliers" (c.f. 4.4.4 "Outliers Management"). This correction considerably improves the model: we will henceforth adopt this new corrected variable.

In the next 4 regressions we combine all of the variables in different ways; the sign of "Airport Density" is always not coherent and "Paved Road" is not significant. We substitute also the "Time Sensitiveness" variable with a Dummy Variable: the coefficient is always significant and the sign doesn't change.

In regression (15) we adjust the "Airport Density" variable excluding the Swiss values. In the correlation analysis we noted that there is distortion caused by Switzerland: the country has the highest airports' density, but the probability of selecting air transport mode is quite low. Although the R^2-Adjusted improves, we get a coherent but not significant coefficient.

In regression (16) we adjust the "Time Difference" variable for Switzerland and Russia introducing time correction that we described above (c.f. 4.4.3 "Independent Variables"). Since the model is better and the coefficient is significant, we will henceforth adopt this corrected variable.

So far, the results are modest. Then, we try to better the model by substituting the dependent variable with "Yspeed%" and "Youtliers%" (Table 32).

DEPENDENT VARIABLE: "Youtliers%"		
REGRESSOR		(1)
Value/Weight Outliers	Coef.	$2*10^{-8}$
	SE	$5.21*10^{-9}$
	t	3.85
Airport Density	Coef.	-0.0013746
	SE	0.0003911
	t	-3.51
GNI/capita	Coef.	$1.78*10^{-6}$
	SE	$4.21*10^{-7}$
	t	4.22
Delta Fares	Coef.	0.0000705
	SE	0.0000498
	t	1.42
Time Difference	Coef.	0.0029317
	SE	0.0005092
	t	5.76
Time Sensitiveness	Coef.	0.1745078
	SE	0.0123779
	t	14.1
Density	Coef.	0.0145774
	SE	0.001663
	t	8.77
R-Squared Adjusted		0.1745
Root MSE		0.19652

DEPENDENT VARIABLE: "Yspeed%"		
REGRESSOR		(1)
Value/Weight	Coef	$4.86*10^{-8}$
	SE	$7.3*10^{-9}$
	t	6.65
Airport Density	Coef.	0.0100356
	SE	0.0004458
	t	22.51
GDP	Coef.	$-7.91*10^{-9}$
	SE	$1.27*10^{-9}$
	t	-6.63
Delta Fares	Coef.	-0.0002446
	SE	0.0000472
	t	-5.18
Time Difference	Coef.	-0.0030579
	SE	0.0005759
	t	-5.31
R-Squared Adjusted		0.2759
Root MSE		0.30348

Table 32 - Linear Regressions with Youtliers% and Yspeed%

Although when we consider "Y%speed" the adjusted-R^2 is higher, many coefficients are not coherent (GDP and time difference are negative). Hence, we reject this option.

On the other hand, the model with "Youtliers%" is consistent and the adjusted-R^2 is better. However some problems persist like the negative value for "Airport Density" and the little significance of "Delta Fares".

Here, we make explicit the best multiple linear regression obtained:

$$\begin{aligned} \boldsymbol{Youtliers}\% = {} & \boldsymbol{\beta_0} + \boldsymbol{\beta_1}(\boldsymbol{V/W\ outliers}) + \boldsymbol{\beta_2}\ \boldsymbol{AirportDensity} \\ & + \boldsymbol{\beta_3}\ \boldsymbol{GNI/capita} + \boldsymbol{\beta_4}\ \boldsymbol{Delta\ Fares} + \boldsymbol{\beta_5}\ \boldsymbol{Time\ Difference} \\ & + \boldsymbol{\beta_6}\ \boldsymbol{Time\ Sensitiveness} + \boldsymbol{\beta_7}\ \boldsymbol{Density} \end{aligned}$$

In general, the multiple linear regression doesn't seem to fit the context properly. We will try other adjustments to improve the model.

5.2.2 STEP (2): Non Linear Regressions

The first regression improvement is to try to correct all of the variables by introducing their most representative function. It's possible that a non-linear function (e.g. quadratic, cubic, logarithmic, etc…) interpolates data better than a linear one.

To accomplish this objective we decide to proceed by "trial and error": for each variable we test different functions and we select the most suitable representation on the basis of the results.

The graphical representation of correlations (c.f. 5.1 "Correlation Analysis"), helps us to imagine "a priori" the best functions.

In general we test and compare three main non-linear functions:

- Logarithmic function → $Y= \beta_0 + \beta_1 \mathbf{ln\ X_1}$
- Quadratic function → $Y= \beta_0 + \beta_1 \mathbf{X_1} + \beta_2 \mathbf{X_1^2}$
- Cubic function → $Y= \beta_0 + \beta_1 \mathbf{X_1} + \beta_2 \mathbf{X_1^2} + \beta_3 \mathbf{X_1^3}$

The introduction of these non-linear functions doesn't change the OLS estimation. In case of logarithmic function, the coefficient interpretation is slightly different: the coefficient doesn't indicate the variation of Y according to a variable's unitary change, but it expresses the variation of Y according to a unitary percentage variation of the variable.

Here, for each independent variable we report a table summarizing the results of the different functions.

DISTANCE

DEPENDENT VARIABLE "Youtliers%"					
REGRESSOR		(1)	(2)	(3)	(4)
Distance	Coef.	.0011055		.0024848	.0018488
	SE	.0000877		.0002628	.0009736
	t	12.60		9.46	1.90
lnDistance	Coef.		5.12409		
	SE		.3951889		
	t		12.97		
SqDistance	Coef.			-9.01e-08	1.04e-08
	SE			1.62e-08	1.49e-07
	t			-5.56	0.07
CbDisance	Coef.				-4.03e-12
	SE				5.94e-12
	t				-0.68
R-Squared Adj.		0.0461	0.0486	0.0547	0.0546

Table 33 - "Distance" Non-Linear Functions

The tests show that both the logarithmic and the quadratic function are consistent. However, the logarithmic coefficient is more significant and it's much simpler to work with one coefficient only. Hence, we select the logarithmic function.

DENSITY

DEPENDENT VARIABLE "Youtliers%"					
REGRESSOR		(1)	(2)	(3)	(4)
Density	Coef.	1.028775		-3.333729	.9268043
	SE	.1404651		.3117821	1.090875
	t	7.32		-10.69	0.85
lnDensity	Coef.		1.119763		
	SE		.4580524		
	t		2.44		
SqDensity	Coef.			.3584405	-.364089
	SE			.0231442	.1788282
	t			15.49	-2.04
CbDensity	Coef.				.0273212
	SE				.0067056
	t				4.07
R-Squared Adj.		0.0225	0.0022	0.1151	0.1211

Table 34 - "Density" Non-Linear Functions

For the "density" we select the quadratic function.

AIRPORT DENSITY

DEPENDENT VARIABLE "Youtliers%"				
REGRESSOR		(1)	(2)	(3)
Airport Density	Coef.	-.0813866	-.6507255	.7943146
	SE	.0263604	.2948586	.1059261
	t	-3.09	-2.21	7.50
SqAirport Density	Coef.		.0855787	-.0186343
	SE		.0199757	.0021847
	t		4.28	-8.53
CbAirport Density	Coef.		-.0015596	
	SE		.0002972	
	t		-5.25	
R-Squared Adj.		0.0026	0.0319	0.0240

Table 35 - "Airports Density" Non-Linear Functions

In this case the best solution is the quadratic function.

PAVED ROADS

DEPENDENT VARIABLE "Youtlier%"					
REGRESSOR		(1)	(2)	(3)	(4)
Paved Road	Coef.	-4.87417		-17.84462	18.69511
	SE	1.478607		12.67696	5.351877
	t	-3.30		-1.41	3.49
lnPaved Road	Coef.		-.8340494		
	SE		.5045093		
	t		-1.65		
SqPaved Road	Coef.			65.26398	-21.7598
	SE			27.78535	4.749876
	t			2.35	-4.58
CbPaved Road	Coef.			-55.48446	
	SE			17.45529	
	t			-3.18	
R-Squared Adj.		0.0030	0.0005	0.0118	0.0091

Table 36 - "Paved Roads" Non-Linear Function

Also in this case the quadratic function seems to be the most suitable.

GNI/CAPITA

DEPENDENT VARIABLE "Youtliers%"					
REGRESSOR		(1)	(2)	(3)	(4)
GNI/capita	Coef.	.0000966		-.002293	.0004112
	SE	.0000225		.0002801	.0001119
	t	4.29		-8.19	3.67
lnGNI/capita	Coef.		1.566358		
	SE		.4157252		
	t		3.77		
SqGNI/capita	Coef.			1.19e-07	-6.07e-09
	SE			1.21e-08	2.11e-09
	t			9.84	1.072686
CbGNI/capita	Coef.			-1.58e-12	
	SE			1.50e-13	
	t			-10.50	
R-Squared Adj		0.0053	0.0040	0.0396	0.0075

Table 37 - "GNI/capita" Non-Linear Functions

The best functional representation of GNI/capita is the cubic one.

TIME DIFFERENCE

DEPENDENT VARIABLE "Youtliers%"				
REGRESSOR		(1)	(2)	(3)
Time Difference	Coef.	.3555983	1.833142	.8210208
	SE	.031912	.2899456	.1319523
	t	11.14	6.32	6.22
SqTime Difference	Coef.		-.1155313	-.0139801
	SE		.0262019	.0038463
	t		-4.41	-3.63
CbTime Difference	Coef.		.002275	
	SE		.0005807	
	t		3.92	
R-Squared Adj		0.0363	0.0441	0.0399

Table 38 - "Time Difference" Non-Linear Functions

The best function for "Time Difference" is the linear one.

FARES DIFFERENCE 1

DEPENDENT VARIABLE "Youtliers%"					
REGRESSOR		(1)	(2)	(3)	(4)
Fares Difference	Coef.	.0100264		.057565	.0428344
	SE	.0024107		.0645532	.0131002
	t	4.16		0.89	3.27
lnFares Difference	Coef.		5.062734		
	SE		1.094204		
	t		4.63		
SqFares Difference	Coef.			-.0000627	-.0000326
	SE			.00013	.0000128
	t			-0.48	-2.55
CbFares Difference	Coef.			1.91e-08	
	SE			8.18e-08	
	t			0.23	
R-Squared Adj.		0.0050	0.0062	0.0063	0.0066

Table 39 - "Fares Difference (1)" Non-Linear Functions

In this case we select the logarithmic function.

FARES DIFFERENCE 2

DEPENDENT VARIABLE "Youtliers%"					
REGRESSOR		(1)	(2)	(3)	(4)
Fares Difference	Coef.	.0020701		-.0031121	-.0066563
	SE	.0002552		.0005602	.0012556
	t	8.11		-5.55	-5.30
lnFares Difference	Coef.		3.225884		
	SE		.6813757		
	t		4.73		
SqFares Difference	Coef.			5.66e-07	1.34e-06
	SE			5.48e-08	2.53e-07
	t			10.32	5.32
CbFares Difference	Coef.				-3.79e-11
	SE				1.20e-11
	t				-3.15
R-Squared Adj.		0.0304	0.0103	0.0776	0.0816

Table 40 - "Fares Difference (2)" Non Linear Functions

All the functions seem to be suitable. However, we select the square function because it has a good R^2 and all the coefficients are significant.

VALUE/WEIGHT(outliers)

DEPENDENT VARIABLE "Youtliers%"					
REGRESSOR		(1)	(2)	(3)	(4)
V/W	Coef.	4.22e-06		.0000494	.0000251
	SE	5.08e-07		3.41e-06	1.74e-06
	t	8.30		14.48	14.44
lnVW	Coef.		6.795865		
	SE		.1925655		
	t		35.29		
SqVW	Coef.			-5.16e-12	-1.23e-12
	SE			4.86e-13	9.85e-14
	t			-10.62	-12.54
CbVW	Coef.			1.35e-19	
	SE			1.64e-20	
	t			8.25	
R-Squared Adj		0.0203	0.2757	0.0838	0.0650

Table 41 - "Value/Weight" Non Linear Functions

Also for the "Value/Weight", the logarithmic representation is the most suitable.

Multiple Non-Linear Regression

Once we have established the best way to represent the variables' functions, we can test them in the complete regression.

We develop two different types of regression; in (1) and (2) we consider "Distance" neglecting "Time Difference" and "Fares Difference" (due to co-linearity's distortions); in (3) and (4) we substitute "Distance" with the other two explicit variables. In general we get a substantial improvement from the multiple linear model (Adj-R^2 changes from 17 to 28/29).

In order to get more legible coefficients, we multiply the dependent variable "Youtliers%" by 100. We report the results in the next table (Table 42).

In the first regression all of the coefficients are coherent except for GNI/capita. In regression (2) we neglect GNI/capita and we obtain almost the same efficacy (Adj R^2 doesn't change substantially); this means that GNI is just a control variable that doesn't contribute to explain strongly the data: for this reason we

exclude it at the moment. By substituting "Distance" with "Delta Fares" and "Time Difference" (in regression n.3) we don't get improvements and the "Fare Difference" coefficient is not significant. Substituting "Fare Difference" with "Fare Difference 2" (c.f. 4.4.3 "Independent Variables"), we get a discrete improvement: Adj-R^2 improves and the coefficients become significant.

REGRESSOR		(1)	(2)	(3)	(4)
lnVW	Coef.	4.161175	4.202095	4.323995	5.060192
	SE	.3061707	.3053059	.3056627	.3136323
	t	13.59	13.76	14.15	16.13
lnDistance	Coef.	3.555313	4.252818		
	SE	.8204926	.6841847		
	t	4.33	6.22		
Airport Density	Coef.	.2994444	.2887809	.286447	.2732586
	SE	.1262972	.1232246	.128332	.1316794
	t	2.37	2.34	2.23	2.08
SqAirport Density	Coef.	-.0046452	-.0044543	-.0080838	-.0076928
	SE	.0028604	.0027968	.0026956	.0027576
	t	-1.62	-1.59	-3.00	-2.79
GNI/capita	Coef.	-.0007129			
	SE	.0003546			
	t	-2.01			
SqGNI/capita	Coef.	3.40e-08			
	SE	1.61e-08			
	t	2.10			
CbGNI/capita	Coef.	-4.39e-13			
	SE	2.07e-13			
	t	-2.12			
Time Sensitiveness	Coef.	5.779185	5.722807	5.530932	5.483393
	SE	1.345703	1.345419	1.35249	1.423713
	t	4.29	4.25	4.09	3.85
Density	Coef.	-1.306405	-1.306729	-1.2934	
	SE	.3521493	.3522108	.3538113	
	t	-3.71	-3.71	-3.66	
SqDensity	Coef.	.1635654	.1628103	.1599028	
	SE	.0251466	.0251534	.0252581	
	t	6.50	6.47	6.33	
lnFares	Coef.			-2.173472	
	SE			1.876332	
	t			-1.16	
Time Difference	Coef.			.2135086	.1592547
	SE			.0482311	.0445064
	t			4.43	3.58
Fares2	Coef.				-.0024355
	SE				.000614
	t				-3.97
SqFares2					3.49E-07
					5.60E-08
					6.23
R-Squared Adjusted		**0.289**	**0.2884**	**0.282**	**0.2903**
Root MSE		18.246	18.246	18.327	18.668

Table 42 - Multiple Non-Linear Regressions

5.2.3 STEP (3): Regressions with Variables Interdependencies

Many times, the effect of the variables on the dependent variable (Y) varies according to other explanatory values.

Let's think about products time difference (from airplanes to ships): it's unlikely that probability of using airplanes varies with the "transportation time difference" in the same way for any product. If a product is very time-sensible, a higher time slack will result in a strong effect on transportation patterns. On the contrary for a product that doesn't have any time constraint, it can even take one year to transport it, but the customers is willing to wait for it so that there's actually no effect on the airplanes usage. We should take into consideration these kinds of products interdependencies to better explain the dependent variables, and to better represent coefficients and effects.

It's possible to study the synergic effect while putting in the regression model a new indicator, built as the product between the two explanatory variables to be analyzed,.

The structure of the equations will be as follows:

$$Y=\beta_0+\beta_1 X_1+\beta_2 X_2+\beta_3 (X_1 * X_2)$$

Where:

- β_1 represents the effect of X_1, if $X_2=0$
- $(\beta_1 + \beta_3 X_2)$ represents the effect of X_1, given a not-null X_2.

Interdependencies Analysis: Testing Intuitive Interdependencies

The most of the variables that we are taking into consideration are linked in some ways. We start now considering pairs of variables for which we can easily suppose there is interaction effect.

DENSITY → UNITARY VOLUME

The variable "Density" should explain the impact of density changes over the dependent variable Y. We discovered that the best relationship is quadratic, but we didn't consider unitary volume constraint.

A very large unitary volume could undermine the possibility to transport by air (volume constraint) or at least to transport by using the traditional pallet (80 cm*120 cm *160 cm). We generate a new indicator called "bulky": a binary variable whose value is "0" if unitary volumes shouldn't be a problem (for the most of the products within a category); the value is "1" if unitary volume can lower air transportation (for many of the products in a given category).

The variable "Density" and the variable "Bulky" seem to be highly interdependent. Density can't have the same impact on transportation patterns, if products have volume constraints.

We try to understand the combined effect through a new variable in the multiple regressions (Table 43) the first regression with the "Distance" factor the second one with "Time Difference" and "Fare Difference" instead of "Distance").

REGRESSOR		(1)	(2)	(3)
ln Value/Weight	Coef.	4.23	4.35	5.04
	SE	0.3	0.3	0.31
	t	14.00	14.39	16.08
ln Distance	Coef.	4.21		
	SE	0.68		
	t	6.22		
Airport Density	Coef.	0.29	0.28	0.27
	SE	0.12	0.13	0.13
	t	2.34	2.23	2.06
Sq Airport density	Coef.	-0.0044	-0.008	-0.007
	SE	0.0028	0.0027	0.0027
	t	-1.59	-3.01	-2.79
Time Sensitiveness	Coef.	6.13	5.94	6.24
	SE	1.33	1.34	1.42
	t	4.6	4.44	4.41
Density	Coef.	-1.09	-1.08	
	SE	0.43	0.43	
	t	-2.56	-2.50	
Sq Density	Coef.	0.165	0.162	
	SE	0.028	0.028	
	t	5.97	5.83	
Bulky	Coef.	8.82	8.83	11.42
	SE	1.43	1.43	2.5
	t	6.19	6.17	4.55
Bulky*Density	**Coef.**	**-0.177**	**-0.178**	
	SE	**0.031**	**0.031**	
	t	**-5.72**	**-5.72**	
Time Difference	Coef.		0.21	0.16
	SE		0.05	0.044
	t		4.42	3.63
ln Fares Difference (1)	Coef.		-2.18	
	SE		1.86	
	t		-1.18	
Fares (2)	Coef.			-0.0015
	SE			0.0008
	t			-1.86
Sq Fares (2)	Coef.			3.14e-07
	SE			6.53e-08
	t			4.80
Bulky*Fares2	Coef.			**-0.178**
	SE			**0.0015**
	t			**-2**
Sq Bulky*Fares2	Coef.			**-9.65e-09**
	SE			**1.68e-07**
	t			**-0.06**
R-Squared Adjusted		**0.3043**	**0.2979**	**0.3049**
Root MSE		18.042	18.123	18.476

Table 43 - Multiple Regressions with "Bulky" and "Density" Interdependences

The multiple regression, in terms of R^2 , is better here with the interdependency factor (regression (1), table 43). All of the coefficients make sense. The sign of the new factor introduced is negative, that means: for products considered

"Bulky" (volume constraints), the probability of using airplanes varies less with density (than for small products).

The coefficient of "Bulky" is positive. It seems that air transportation would be easier for bulky products. However, the coefficient represents the effect in absence of density (density = 0) and it's impossible to have products with "0" density. The real effect for light products can be evaluated solving the equation while putting the smallest density value.

The "interaction variable" here reported is actually sum of two factors (since we set a quadratic function to describe density):

Int Density-Volume= [(bulky)*(density)] + [(bulky)*(Density)^2]

By substituting distance variable with time and fare slacks (regression (2), table 43), regression shows very similar results: time and fare differences are co-linear with distance.

We would expect better results here, since the two variables should mitigate for omitted variables errors. On the contrary, the R^2 indicator lowers, and the "ln fare difference" t-value is less than 1.96. Probably, we didn't use a very precise fare indicator.

We can better analyze the concurrent impact of "Times Difference" and "Fares1" through a simple regression (table 44). We would expect a positive coefficient for time difference: if there is time advantage, then airplanes are more convenient. On the contrary, we would expect negative sign for fares that vary with distance: if it's more expensive to carry something, then airplanes should be less convenient.

REGRESSOR		
Ln Fares1	Coef	-2.69
	SE	1.312
	t	-2.05
Time Difference	Coef	**0.40**
	SE	0.039
	t	10.31
R-Squared Adjusted	**0.0372**	
Root MSE	20.942	

Table 44 - Simple Linear Regression with "Fares(1)" and "Time Difference"

The “Fares1” coefficient is high and negative, but it’s much less statistically significant. Standard error is high and we can’t refuse H_0 with 99% confidence. We have already noticed before that in all of the multiple regressions in which we substituted distance with “Time Difference” and “Fare1”, the fare coefficient is “weird” (we said that it could be the result of not precise measures). However, looking at the last regression it appears clear that impact of distance is statistically consistent almost only from a time perspective. The conclusion is that if someone is thinking about using airplanes, it’s not a matter of distance itself (there in no causal effect) and it’s not even a matter of costs: actually the point is not how much expensive is to carry something, since probably that delivery is considered urgent.

We have already introduced the “Fares2” indicator. This represents fare structure varying distance and **density** parameters: we could substitute density (and “Fare1”) with “Fares2” in the regressions (Regression (3), table 43). So, costs don’t matter from a distance perspective. But, as already said in the technical chapter, costs vary not only with distance (oil, resources consumption). Especially freight costs vary with density of products transported, since sea fares changes with volume of freight, while air fares are usually more weights sensitive (weight is a more restrictive constraint).

Being “Fares2” linked with “Density”, we expect interaction with the variable “Bulky”.

We represented in table both of the interaction terms ("Fares2" is a quadratic function) and the coefficients are negative. Interaction coefficient lowers the impact of higher tariffs. Results are consistent with previous regressions: "Density" and "Fares2" have similar interaction with the variable "Bulky".

In conclusion, we can substitute "Density" with "Fares2" in regressions, while we should avoid to include "Fares1" to replicate effect of distance.

TIME DIFFERENCE or DISTANCE → TIMELINESS

The second interaction we consider is between "Time Difference" and "Timeliness". If time advantage by air is bigger, there should be higher probability of using airplanes. But it's clear; it's not just a matter of time: if the product to be delivered is very time sensitive there will be a much stronger impact. We could analyze many interdependencies among variables and time difference, but this is the easier relation to be forecasted: "Time difference" and "Timeliness" are independent variables, but the relative effect should be strongly related.

We report in regression (1) (table 45) the obtained results. The interaction between "Time Difference" and "Timeliness" is strong (0.46) and consistent (t-value = 4.76): if a product is time-sensible, then the effect of longer time slacks is higher and higher and it increases the probability of using airplanes.

REGRESSOR		(1)	(2)
ln Value/Weight	Coef.	4.43	4.33
	SE	0.3	0.3
	t	14.69	14.49
ln Distance	Coef.		-6.06
	SE		1.6
	t		-3.79
Time Difference	Coef.	**-0.35**	
	SE	0.13	
	t	-2.74	
Airport Density	Coef.	0.28	0.28
	SE	0.13	0.12
	t	2.21	2.37
Sq Airport density	Coef.	-0.008	-0.0045
	SE	0.003	0.0027
	t	-3	-1.63
Time Sensitiveness	Coef.	**-0.55**	**-65.25**
	SE	1.9	10.19
	t	-0.29	-6.4
Density	Coef.	-1.06	-1.07
	SE	0.43	0.42
	t	-2.49	-2.54
Sq Density	Coef.	0.16	0.16
	SE	0.028	0.027
	t	5.80	5.94
Bulky	Coef.	8.85	8.84
	SE	1.42	1.4
	t	6.22	6.28
Bulky*Density	**Coef.**	-0.18	-0.18
	SE	0.03	0.03
	t	-5.8	-5.85
Int Time-Timeliness	Coef.	**0.46**	
	SE	0.096	
	t	4.76	
Int Distance-Timeliness	Coef.		**8.4**
	SE		1.19
	t		7.06
R-Squared Adjusted		**0.3055**	**0.3242**
Root MSE		18.026	17.825

Table 45 - Multiple Regression with "Time Difference/Distance" and "Timeliness" Interdependency (a)

The coefficient of “Timeliness” is negative. This is for sure counter-intuitive because it seems that it’s less likely to use airplanes for more time-sensible products. However, coefficient itself explains only the effect in the hypothesis of having null time slack: if airplanes aren’t faster than alternative means of transport, why should air probability rise for time-sensible products? P-value for timeliness is 0.77. It means that there is 77% of probability for the coefficient to be “0”.

"Time Difference" coefficient is negative: for products that are totally not sensible to time, the higher the time slack, the lower the probability of using airplanes. It actually makes sense: higher time difference implies higher distance and costs to be sustained. The coefficient is consistent because P-value is 0.006 (6% of probability for the effect to be null).

In the context of this model, "Time Difference" is a measure very similar to distance: we will test now the dependence between "time sensitiveness" and "distance". The results are reported in regression (2) in table 45.

The interaction factor here analyzed is very strong and consistent: the interdependency must be considered in the model.

As for the previous regressions, R^2 is higher when we consider distance instead of the two explanatory variables of time and fares. However, there is a weird coefficient.

After having analyzed the previous regression coefficients, we would expect null coefficient for timeliness here (if there is no time difference why should Y be different for time sensible products?). On the contrary, we find a very strong (-65.25) and consistent (-6.4) negative ß "timeliness" (regression (2), table 45). **It means that if distance is "0" then time sensible products are much less likely to be delivered by air.** The result is counterintuitive but it's actually correct considering that airplanes have much higher "fixed times" than trucks, so that for (unrealistic) distances close to "0", airplanes aren't the fastest mean of transport: if you have to deliver something in your neighborhood, airplanes are out of context!

The coefficient we get here is not very significant for our analysis because in international trade it doesn't make any sense to have null distance. The problem is that, as already said above, distance isn't actually a direct "explanatory factor" for airplanes usage: in the reality it's not a matter of distance itself, it's a matter

of time and costs advantage. However, we could approximate the real impact of timeliness while putting the closest country distance (Switzerland) equal to "0" and adjusting consequently the other values.

Moreover, our arbitrary timeliness indicator could introduce error since values range from 1 to 3 for the most of the categories: a unitary change for timeliness implies very strong effect and if the evaluation hasn't been very precise, it can result in big distortion.

Finally, coefficients must be "read" in a proper way. We have logarithms of distances, so that coefficients are not the effect of a unitary change in variables: "ln distance" coefficient is -6.06 but it doesn't mean that for any added km there's a 6 points lower Y. Even timeliness effects, combined with distance here, must be interpreted.

We can verify our assumptions by:

- **A.** Considering linear relationship (instead of logarithmic) for distance.
- **B.** Changing "Timeliness" in a binary variable to reduce the arbitrariness of the evaluation: we set 1 for values higher than 1.8 while obviously setting 0 for lower values.
- **C.** Adjusting "Distance" by putting the closest country distance (Switzerland) equal to "0" and consequently computing other values.

In the following table (table 46) we report the multiple regressions' results that we obtained in the different tests:

REGRESSOR		(1)	(2)	(3)
ln Value/Weight	Coef.	4.35	4.32	4.16
	SE	0.3	0.28	0.26
	t	14.49	15.45	15.94
Distance	Coef.	**-0.001**		
	SE	0.0003		
	t	-3.81		
Ln Distance	Coef.			**3.17**
	SE			0.6
	t			5.22
Time Difference	Coef.		**0.157**	
	SE		0.048	
	t		3.24	
Airport Density	Coef.	0.44	0.28	0.24
	SE	0.12	0.125	0.1
	t	3.75	2.25	2.25
Sq Airport density	Coef.	-0.009	-0.008	-0.004
	SE	0.002	0.0026	0.0024
	t	-4.08	-3.04	-1.46
Time Sensitiveness	Coef.	**-4.24**		
	SE	2.19		
	t	-1.94		
Time Sensitiveness Dummy	Coef.		**2.6**	**-76.23**
	SE		2.22	11.97
	t		1.17	-6.36
Density	Coef.	-1.08	-1.14	-1.28
	SE	0.426	0.42	0.39
	t	-2.53	-2.72	-3.23
Sq Density	Coef.	0.16	0.173	0.18
	SE	0.027	0.027	0.026
	t	5.90	6.35	7.23
Bulky	Coef.	8.84	8.15	8.32
	SE	1.42	1.41	1.34
	t	6.24	5.76	6.23
Bulky*Density	Coef.	-0.18	-0.17	-0.18
	SE	0.03	0.03	0.027
	t	-5.81	-5.55	-6.77
Int Time-Timeliness	Coef.		**0.568**	
	SE		0.118	
	t		4.81	
Int Distance-Timeliness	Coef.	**0.0015**		
	SE	0.00027		
	t	5.88		
Int ln Distance - Timeliness	Coef.			**10.29**
	SE			1.4
	t			7.33
R-Squared Adjusted		**0.3128**	**0.3156**	**0.3261**
Root MSE		17.931	17.894	17.019

Table 46 - Multiple Regression with "Time Difference/Distance" and "Timeliness" Interdependency (b)

A) LINEAR RELATIONSHIP TO DISTANCE

The regression coefficients are reported in "regression (1)" in the table 46.

As expected by considering "distance" as a linear function, signs are consistent but "sizes" change a lot: one added km implies a 0.001 reduction of Y.

B) DUMMY VARIABLE FOR TIMELINESS

Then we test the regression, introducing the dummy variable for "timeliness" combined with "time difference" or "ln distance". The results are reported in "regression (2)" in table 46.

"Time Difference" coefficient changes sign. This results from the new timeliness definition. The most of the categories have "0" timeliness and the coefficient shows the average effect for all of these products. For "normal" products, an increase in time difference results in higher probability of using planes because of bigger time advantages (in the previous regressions the 0 timeliness represented unrealistic products since the minimum value was 1, while now the null effect represents the most of the products in the database).

Even the "Time Sensitiveness" coefficient changes and it's now positive. This shows the effect of passing from 0 to 1 in timeliness, (we are now considering very time-sensible products) under the assumption of having null time slack. As discussed before, we expected a null coefficient (if airplanes aren't faster than alternative means of transport, why should air probability rise for time-sensible products?).

However, the low "t-value" (1.17) allows us not refusing the hypotheses of having null impact.

Now we substitute "Time Difference" with "lnDistance" and we test the interdependence with the new dummy variable "Time Sensitiveness".

Firstly, we test the interdependence in a Simple Regression. We report the results in the table below (table 47).

REGRESSOR		
Ln Distance	Coef	**3.89**
	SE	0.43
	t	9.11
Time Sensitiveness	Coef	**-65.86**
	SE	10.2
	t	-6.45
int Distance-Timeliness	Coef	9.96
	SE	1.199
	t	8.31
R-Squared Adjusted	0.1545	
Root MSE	19.563	

Table 47 - Simple Regression with "Distance-Timeliness" Interdependency

The interaction is strong in this case too: “Int Distance-Timeliness” coefficient is 9.96 and it’s statistically significant (t =8.31).

Then we test this significant interdependence in the complete Multiple Regression. The results are reported in “Regression (3)” in table 46.

The interaction factor has high and significant coefficient: distance effect is much higher for time sensible products.

The distance coefficient is positive and statistically significant: even for “normal” products, the higher the distance the higher the airplanes usage.

The timeliness coefficient is negative and statistically significant: under the assumption of having null distance, airplane is much less probable for very time-sensible products. As already said above, the coefficient looks quite misleading since we would expect higher probability for time sensible products towards any destination. However, there can be no trade with null distance: we will show how results change considering the zero-distance for the closest country Switzerland.

C) CORRECTED DISTANCE "DISTANCE 0"

Interaction "Distance 0" and Time Sensitiveness:

REGRESSOR		
Ln Distance0	Coef	**1.04**
	SE	0.153
	t	6.77
Time Sensitiveness	Coef	**-3.49**
	SE	3.53
	t	-0.99
INT Distance0 -Timeliness	Coef	2.84
	SE	0.43
	t	6.57
R-Squared Adjusted	0.1265	
Root MSE	19.885	

Table 48 - Simple Regresion with "Distance 0- Time Sensitiveness" Interdependency

The "Time Sensitiveness" coefficient is still negative, but statistically not very relevant: the P-Value is 0.32. It means with a 32% probability, the effect is null. Moreover, time difference considered for Switzerland is -0.5 days. Airplane is a little slower than trucks if Switzerland is the partner. So, the result is really coherent.

Interdependencies Analysis: Testing Other Less Intuitive Interdependences

We have analyzed the most logical relationships among variables. Now, we will go through some other tests to verify interdependencies among the most relevant variables.

"Distance" (or "Time Difference") and "Value-Weight" are the factors with the strongest impact over Y. All of the other variables effect could depend on them. We will try to understand if variables interdependencies with distance (or time difference + fares difference) and value/weight are important enough to be included in model.

Focus on "Distance"

DISTANCE → FARES / DISTANCE → TIME DIFFERENCE

There is imperfect co-linearity between: "Distance" and "Fares Difference", "Distance" and "Time Difference", "Fares Difference" and "Time Difference". From a logical perspective, it doesn't make sense to introduce interdependencies factors among them. That's the reason why we expect distortion in the following regressions.

REGRESSOR		(1)	(2)	(3)
ln Fares1	Coef.	**0.78**		**-29.6**
	SE	1.63		8.47
	t	0.48		-3.49
Time Difference	Coef.	4.89	**-0.65**	
	SE	1.27	0.56	
	t	3.85	-1.16	
ln Distance	Coef.		4.08	-9.65
	SE		0.6	5.95
	t		6.75	-1.62
INT Fares1-Time	Coef.	-0.712		
	SE	0.201		
	t	-3.53		
INT Distance – Time	Coef.		0.08	
	SE		0.06	
	t		1.35	
INT Distance-Fares1	Coef			2.77
	SE			1
	t			2.76
R-Squared Adjusted		**0.0406**	**0.0499**	**0.0567**
Root MSE		20.905	20.80	20.73

Table 49 – Simple Regressions with "Fares/Time Difference" and "Distance"

In all the regressions, R-squared is almost the same considering only one factor at the time, this to confirm the distortion from co-linearity.

DISTANCE or TIME DIFFERENCE→ DENSITY

Density of products can influence transportation patterns. There could be different "density impact" over Y, if transportation implies a long or short distance to be covered.

REGRESSOR		(1)	(2)	(3)
ln Distance	Coef.	4.42	4.08	
	SE	0.43	0.6	
	t	10.8	6.75	
Time difference	Coef.			0.3217
	SE			0.03743
	t			8.6
Density	Coef.	-3.35	-0.65	-3.354
	SE	0.42	0.56	0.4265
	t	-7.91	-1.16	-7.86
Sq Density	Coef.	0.37	0.08	0.3642
	SE	0.026	0.06	0.028
	t	14.05	1.35	12.69
Bulky	Coef.	8.59	8.58	8.55
	SE	1.48	1.48	1.48
	t	5.81	5.80	5.75
Int Bulky-Density	Coef.	**-0.14**		-0.142
	SE	0.03		0.03
	t	-4.67		-4.66
Int Density-Distance	Coef.		0.017	
	SE		0.0098	
	t		1.75	
Int Density-Time	Coef.			0.00049
	SE			0.00079
	t			0.62
R-Squared Adjusted		**0.17**	**0.1708**	**0.1611**
Root MSE		18.888	1.879	18.988

Table 50 - Simple Regressions with "Distance/Time Difference" and "Density"

We consider here the interaction factor as a single variable even if it can be actually divided in two terms being density a quadratic function.

This interaction factor is not very significant for both the “Distance” and the “Time Difference” variable (t-value are 1.75 and 0.62). Moreover, the R-Squared is almost the same in regressions with or without combined effect. We won’t considerate this interaction.

DISTANCE or TIME DIFFERENCE → VALUE/WEIGHT

We consider now the two "strongest" variables, testing interaction between them.

REGRESSOR		(1)	(2)
Ln Distance	Coef	-16.03	3.54
	SE	1.38	0.34
	t	-11.64	10.33
Ln Value/Weight	Coef	-15.8	6.53
	SE	1.54	0.19
	t	-10.29	34.12
INT Distance-Value/Weight	Coef	2.68	
	SE	0.18	
	t	14.64	
R-Squared Adjusted		**0.34**	**0.298**
Root MSE		17.321	17.878

Table 51 - Simple Regressions with "Distance" and "Value/Weight" Interdependency

The interaction coefficient is positive and statistically very significant: the Value/Weight effect on Y increases with relative distance and vice-versa.
The coefficients of distance and value/weight by themselves are quite misleading. They are negative, while we would logically expect air probability to rise for higher distances such as higher value/weights (as in the "normal regression": regression (2), table 51).

In the reality, the result is coherent with what said above because if distance is null, then airplanes aren't faster than other mean of transportation and trucks are preferable, especially for those goods with high price (value/weight). Moreover, under the unrealistic assumption of having products with null price (value/weight) it doesn't make any sense to transport them on fast and expensive mean of transport, especially for longer routes.

The regression with interaction looks much better in terms of R-squared and it's actually better to consider this combined effect. However, we need to fully understand the impact of the negative coefficients.

Let's try to check for coefficients, considering linear functions (instead of logarithmic), so that coefficients represent impact of unitary changes.

REGRESSOR		
Distance	Coef	0.001
	SE	0.00008
	t	12.60
Value/Weight	Coef	1.37e-07
	SE	7.45e-07
	t	0.18
INT Distance-Value/Weight	Coef	1.54e-09
	SE	2.01e-10
	t	7.66
R-Squared Adjusted	**0.0847**	
Root MSE	20.42	

Table 52 - Simple Regressions with "Distance" (linear) and "Value/Weight"(linear) Interdependency

The interaction is still important, the distance coefficient becomes significantly positive, while value weight coefficient becomes positive with high P-value=0.85 (the effect is very likely to be null). However, R-squared is very low: we already know that linear relationships are not very adept at the context.

Here we test interaction with linear distance and logarithmic value/weight.

REGRESSOR		
Distance	Coef	-0.0034
	SE	0.00033
	t	-10.23
Ln Value/Weight	Coef	3.2
	SE	0.32
	t	09.91
INT Distance-Value/Weight	Coef	0.00056
	SE	0.00004
	t	12.62
R-Squared Adjusted	**0.3263**	
Root MSE	17.529	

Table 53 - Simple Regressions with "Distance" (linear) and "Value/Weight"(logarithmic) Interdependency

The regression has quite high R-Squared, but signs are still inconsistent: we had both of the coefficient negative in the first regression, they were positive with linear functions, they are different in sign now. We have confirmation that the combined effect is strong and positive. **Distance has "more positive" effect for higher value/weight and/or value/weight has "more positive" effect for higher distances.**

Equation is structured as follows:

$Y=ß_0+ß_1X_1+ß_2X_2+ß_3(X_1*X_2)$

In the first regression we had X_1=ln(distance) and X_2=ln(Value/Weight)

*Y=95.08 – 16*ln(distance) – 15.8*ln(value/weight) + 2.68*[ln(valueweight)*ln(distance)]*

*δY/δ (lndistance)= -16 + 2.68*ln(value/weight)*

Y'=0 → ln(value/weight) = 5.97 → value/weight = $e^{5.97}$= 391 €/tons

*Y = 95.08 - 16 X - 15.8*5.97 + 2.68 * 5.97 * X = 0.754 %*

When value/weight is about 391 €/tons, the first derivate (slope of the line) is equal to 0 so that probability of using airplanes doesn't vary with "ln(distance)". For any distance there's a 0.754% probability to transport by air. For smaller "prices" the line is decreasing: higher distances imply smaller airplanes usage.

On the contrary, if we solve the derivate fixing ln(distance) we have:

*δY/δ(lnvalue/weight) = -15.8 + 2.68 *ln(distance)*

Y'=0 → ln(distance) = 15.8/2.68 = 5.89 → distance = $e^{5.89}$ = 361 km

*Y = 95.08 - 16*5.89 - 15.8*X + 2.68 * 5.89 * X = 0.84%*

For transportation 361 Km long, Value/Weight doesn't change probability of using planes. We have a 0.84% for any kind of products' price. For smaller distances we have decreasing line and probability.

Let's try to solve equations referred to the third regression implemented (linear distance; ln Value/Weight).

*Y = -17.63 - 0.0034*distance + 3.2*ln(value/weight) + 0.00056*[distance*ln(value/weight)]*

*δY/δ(distance) = -0.0034 + 0.00056*ln(value/weight)*

Y'=0 → ln(value/weight)=6.07 → value/weight = $e^{6.07}$= 432 €/tons

*Y = -17.63 + 3.2*6.07 = 0*

Or:

*δY/δ(lnvalue/weight)= 3.2 + 0.00056*distance*

Y'=0 → distance = -5714 →We have only positive slopes

Results are different because we have measures of different nature (and different **constant value**). A good test is to substitute given values and control for consistency of the probability in output.

For: distance = 2000 km and V/W = 1000 €/tons→ ln(value/weight)=6.9; ln(Distance)=7.6

Reg1→ Y= 5%

Reg2→ Y= 5,3 %

For: Distance = 10000 km; Value/Weight=10000 €/tons→ ln(distance)= 9.21; ln(v/w) = 9.21

Reg1 → Y = 29.5%

Reg2 → Y = 29.4%

The results are quite consistent; however we will refer to the first regression that has the highest R-Squared.

It's easy to understand that for "cheap" products under 391 €/tons it's less likely to use airplanes if distance is bigger (even the fare difference is bigger). On the contrary, it seems strange that there is negative slope if distances are smaller than 361 Km. However, we have to consider that airplanes are in average slower than trucks if the destination is Switzerland (in average about 600 km far).

Let's calculate the slope referred to average distance of Switzerland. We would expect almost null coefficient since time difference is not very relevant. To evaluate it, we go through a new regression setting a null distance for

Switzerland (so that the ln (value/weight) coefficient represents the actual slope of the line).

REGRESSOR		
Distance0	Coef	-4.63
	SE	0.46
	t	-10.08
Ln Value/Weight	Coef	0.92
	SE	0.49
	t	1.88
INT Distance-Value/Weight	Coef	0.78
	SE	0.06
	t	12.73
R-Squared Adjusted	0.3244	
Root MSE	17.544	

Table 54 – Simple Regression with "Distance 0" and "Value/Wieght" Interdependency

The coefficient is positive even if t-value is not very high. It means that probability of using airplanes to go in Switzerland increases slowly with products price. However, we can't refuse the hypotheses of having null effect.

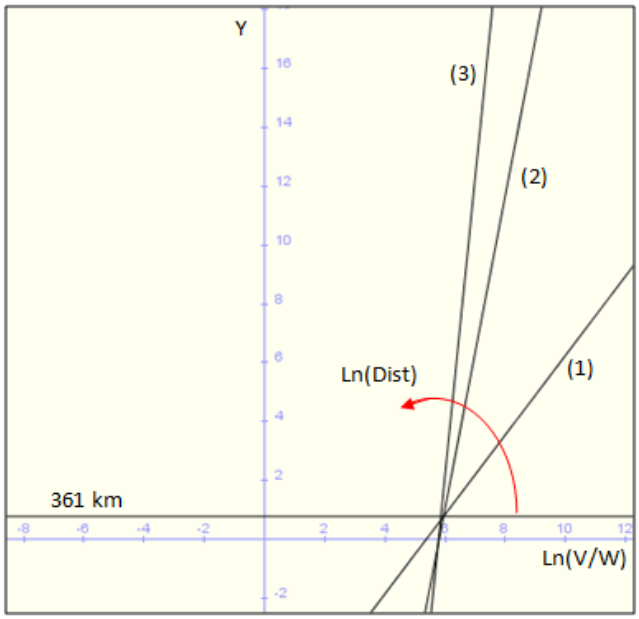


Figure 32 - Derivates' Functions Representation

(1)=Switzerland; (2)=Russia; (3)= Australia.

The interaction between "Distance" and "Value/Weight" is relevant and must be considered in the multiple regressions.

Here the interactions in the multiple regression:

REGRESSOR		(1)
ln Value/Weight	Coef.	-12.06
	SE	1.95
	t	-6.20
Ln Distance	Coef.	**-10.65**
	SE	1.78
	t	-5.99
Airport Density	Coef.	0.27
	SE	0.12
	t	2.28
Sq Airport density	Coef.	-0.004
	SE	0.0026
	t	-1.67
Time Sensitiveness	Coef.	-45.79
	SE	12.82
	t	-3.57
Density	Coef.	-1.14
	SE	0.409
	t	-2.78
Sq Density	Coef.	0.173
	SE	0.0266
	t	6.53
Bulky	Coef.	7.75
	SE	1.38
	t	5.63
Bulky*Density	Coef.	-0.164
	SE	0.03
	t	-5.49
Int Distance- Value/weight	Coef.	1.96
	SE	0.23
	t	8.44
Int ln Distance - Timeliness	Coef.	6.58
	SE	1.5
	t	4.37
R-Squared Adjusted		**0.3528**
Root MSE		17.401

Table 55 - Multiple Regression with "Value/weight –Distance" and "Timeliness-Distance" Interdependency

"Ln Distance" is negative and significant. This refers to products with "0" timeliness and "0" Value/Weight concurrently.

We can replicate effect of distance by using time difference. This should bring improvement to analysis since there is actual casual link between the new variable and Y.

We test interdependency between "Time Difference" and "Value/Weight", expecting consistency with previous results.

REGRESSOR		
Time difference	Coef	-1.165
	SE	0.127
	t	-9.13
Ln Value/Weight	Coef	4.354
	SE	0.2763
	t	15.75
INT Time Difference-Value/Weight	Coef	0.1833
	SE	0.0166
	t	11.3
R-Squared Adjusted	0.3133	
Root MSE	17.687	

Table 56 – Simple Regression with "Time Difference" and "Value/Weight" Interdependency

Results are consistent with expectations. We can replicate effect of distance and in particular effect of distance interdependency with value/weight by using the time difference variable.

Focus on "Value/Weight"

Let's test all of the interdependencies with Value/Weight.

VALUE WEIGHT → TIMELINESS

REGRESSOR		
Timeliness	Coef	-64.99
	SE	5.86
	t	-11.08
Ln Value/Weight	Coef	5.173
	SE	0.215
	t	24.05
INT Timeliness -Value/Weight	Coef	8.6
	SE	0.67
	t	12.74
R-Squared Adjusted	0.3310	
Root MSE	17.401	

Table 57 - Simple Regression with "Timeliness" and "Value/Weight" Interdependency

Interaction is strong and significant. If products are time sensitive, then the impact of value/weight is stronger (and vice-versa). The coefficient of "LnValue/Weight" is positive and significant as expected. For normal product (timeliness=0), air probability is still increasing with value/weight.

The coefficient of the timeliness is strongly negative and significant. This result is quite unexpected, since logically it sounds weird that products are less likely to be transported by air if they have time constraints. The coefficient reflects the slope in the case of having value/weight (a proxy for products price) equal to 0.

It can be interesting to understand what is the "threshold value" that makes the slope equal to "0".

Here the equation to be considered:

$Y= -28.67 – 64.99X_1 + 5.173X_2 + 8.6X_1X_2$

$\delta Y/\delta X_1= -64.99 + 8.6X_2$

$Y'=0 \rightarrow X_2= 7.56 \rightarrow$ Value/Weight = 1919.85

Y = -28.67 + 5.173*7.56 = 10.44%

For value/weight equal to 1919.85 €/ton, the presence of time constraints is irrelevant (while for lower prices, airplanes usage goes down if products are time-sensitive). It seems a paradox since it doesn't make sense to transport slowly products that require speed.

The point is that we can't only consider transportation patterns because this result is probably consequence of particular trade patterns. **If someone needs fast delivery, he/she can use airplanes or he/she can rather trade with closer countries (or produce it by him/herself).**

The categories whose value/weight is lower than 1900 €/ton (and time constrained) are probably traded almost only with close partners, so that the probability of using airplanes becomes lower. Let's think about milk (category "143 milk, cream and fresh"): its value weight is lower than 1900, and it has a restrictive perishability constraint. Milk and creams trade can involve only close partners because air transportation costs are too high if compared with unitary production costs. It's much better to import them from "neighbors" (or better to self-produce it) since the higher production costs are still lower than transportation surplus.

Moreover, the most of products with these features are "food products". In the reality, almost any country is willing to satisfy by itself its own food requirement even if there is no comparative advantage. The fear of wars or crisis pushes the Governments towards protectionist policies, in order to enforce local producers.

We can conclude saying that in average, need for time influences transportation choice only for products more expensive than 1920 €/ton. Under this threshold there's no trade off: the need for time is more likely to influence trade patterns (and only cheap means will be

used) because production savings are lower than fast transportation extra-costs.

VALUE WEIGHT → DENSITY

REGRESSOR		(1)	(2)
Ln Value Weight	Coef	5.4	5.26
	SE	0.409	0.25
	t	13.21	21.04
Density	Coef	-3.9	-1.39
	SE	1.72	0.41
	t	-2.27	-3.42
Sq Density	Coef	0.58	0.17
	SE	0.138	0.026
	t	4.18	6.3
Bulky	Coef	11.06	8.96
	SE	1.478	1.39
	t	7.51	6.46
Int Bulky-Density	Coef	-0.26	-0.19
	SE	0.034	0.028
	t	-7.59	-6.78
Int Density-VW	Coef	0.121	
	SE	0.21	
	t	0.58	
Sq Int Density-VW	Coef	-0.026	
	SE	0.012	
	t	-2.02	
R-Squared Adjusted		**0.2745**	**0.2694**
Root MSE		17.658	17.72

Table 58 - Simple Regression with "Value/Weight" and "Density" Interdependency

Coefficients related to interaction are difficult to be read (and not very significant). Moreover, R-squared is almost the same with or without considering combined effect (0.2745 vs 0.2694). Interdependency doesn't seem important and we won't consider it.

Final multiple regression considering all of the relevant interactions

Now we compute complete regressions by including all of the relevant variables' interdependencies that we pointed out.

We develop three final regressions (table 59): the first one (1) includes "Distance", in the second one (2) we substituted "Distance" with "Time Difference" and in the last one (3) we introduced "Fares2" (substituting "Density").

The first regression has the highest R-Squared (38.18%) and their coefficients are statistically significant. In the second regression we introduce "Time Difference", to substitute "Distance". The interaction with timeliness doesn't have a very significant coefficient (we can't refuse H_0 with a 90% confidence); moreover R-Squared lowers a little (36.16%). Probably there is omitted variable error since we try to "summarize" effect of distance while only introducing a time indicator ("Fares1" was co-linear with time difference and we decided to exclude it). We introduce "Fares2" instead of "Density" in the third regression. This indicator keeps into consideration fares difference according to distance and density concurrently. R-squared lowers a little more and the interaction "Time Difference-Timeliness" keeps being not very significant.

As already said above, coefficient signs seem to be not always coherent (e.g. "Bulky" is positive, "Timeliness" is strongly negative). However, we have to consider them together with the interaction factors.

REGRESSOR		(1)	(2)	(3)
ln Value/Weight	Coef.	-11.77	1.99	2.95
	SE	1.902	0.37	0.4
	t	-6.19	5.37	7.36
Ln Distance	Coef.	-9.536		
	SE	1.74		
	t	-5.48		
Time Difference	Coef.		-0.74	-0.562
	SE		0.15	0.169
	t		-4.78	-3.31
Delta Fares2	Coef.			-0.0017
	SE			0.0006
	t			-2.68
Sq Delta Fares2	Coef.			3.74e-07
	SE			5.54e-08
	t			6.75
Airport Density	Coef.	0.27	0.27	0.24
	SE	0.114	0.119	0.125
	t	2.35	2.30	1.99
Sq Airport density	Coef.	-0.004	-0.0075	-0.0073
	SE	0.0026	0.0025	0.0026
	t	-1.76	-2.98	-2.80
Time Sensitiveness	Coef.	-104.73	-74.74	-76.44
	SE	13.9	8.17	8.98
	t	-7.53	-9.15	-8.51
Density	Coef.	-1.423	-1.44	
	SE	0.401	0.407	
	t	-3.55	-3.54	
Sq Density	Coef.	0.204	0.206	
	SE	0.026	0.026	
	t	7.81	7.74	
Bulky	Coef.	5.92	5.98	4.65
	SE	1.36	1.38	1.3
	t	4.36	4.34	3.56
Bulky*Density	Coef.	-0.13	-0.13	
	SE	0.03	0.03	
	t	-4.58	-4.55	
Int Fares2 - Bulky	Coef.			-2.97e-07
	SE			6.31e-08
	t			-4.71
Int Distance- Value/weight	Coef.	1.82		
	SE	0.23		
	t	8.02		
Int Time Difference-Value/weight	Coef.		0.123	0.095
	SE		0.02	0.02
	t		5.94	4.22
Int Distance - Timeliness	Coef.	4.323		
	SE	1.49		
	t	2.9		
Int Time Difference-Time Sensitiveness	Coef.		0.144	0.18
	SE		0.12	0.022
	t		1.19	1.39
Int Value/Weight - Timeliness	Coef.	9.09	9.63	9.69
	SE	0.93	0.95	1.01
	t	9.75	10.16	9.53
R-Squared Adjusted		**0.3818**	**0.3616**	**0.3597**
Root MSE		17.005	17.282	17.732

Table 59 - Comple Multiple Regressions with Interdependencies (a)

5.2.4 Scenarios Analysis

In the previous step we discovered the relevant role of interdependencies among variables. We can state that the probability of "air shipping" doesn't depend on a single variable but it is concurrently affected by the interactions between different determinants.

For this reason, it's impossible to interpret the single contribution of each variable: the effect on Y of a variable's variation depends also on the value assumed by all of the other interdependent variables. For example, if value/weight increases the effect on "air shipping" probability is different according to the distance to be covered and to the time sensitiveness of the cargo.

So, we decide to implement "scenarios analyses" in order to capture the effect of each single variable making the other variables vary.

In particular, we develop two scenarios analysis: the first one assesses the variation of "Distance", the second one focuses on "Timeliness".

SCENARIOS ANALYSIS 1

	TIME SENSITIVENESS	
VALUE/WEIGHT	Low; 0 (1^)	Low; 1 (2^)
	High; 0 (3^)	High; 1 (4^)

The value on the left in the matrix indicates the value/weight, while the one on right refers to time sensitiveness. We consider:

- V/W = "Low" = 100 €/ton
- V/W= "High" = 2000 €/ton
- Time Sensitiveness = "Low" = 0
- Time Sensitiveness = "High" = 1

Now, we simulate the effect of distance for each scenario by considering a distance increase from 500 km to 6000 km. The basic question is: "how does the "air shipping probability" change if the distance to be covered increases from 500 km to 6000 km?". In simulations we used the better regression model that we developed so far (R^2-Adjusted = 0.3818 → Reg(1), table 59).

ΔY= f (ΔDistance) = f(Distance = 6000 km) – f(Distance=500)

*ΔY= -9.55 * ln (Distance "High") + 4.32 [ln(Distance "High")*Time Sensitiveness] + 1.82 [ln(Distance "High")*ln(W/V)] – {(-9.55) * ln (Distance "Low") + 4.32 [ln(Distance "Low")*Time Sensitiveness] + 1.82 [ln(Distance "Low")*ln(W/V)]}*

Scenario No.1 (Low-0)

Distance "High":

Y= -9.55 * ln(6000) + 4.32 [ln(6000)*0] + 1.82 [ln(6000)*ln(100)] = -10,17 %

Distance "Low":

Y= -9.55 * ln(500) + 4.32 [ln(500)*0] + 1.82 [ln(500)*ln(100)] = -7,25 %

ΔY = -10.17%– (-7,25%) = **-2.92%**

In the first scenario the probability of using plane lowers. The result is intuitive; we are considering a product that is not time sensible, so that even though distance and delivery time are higher, it's possible to take a slower transport mode. Moreover, the low Value/Weight makes "air shipping" costs unsustainable.

Scenario No.2 (Low-1)

Distance "High":

Y= -9.55 * ln(6000) + 4.32 [ln(6000)*1] + 1.82 [ln(6000)*ln(100)] = +27.41%

Distance "Low":

Y= -9.55 * ln(500) + 4.32 [ln(500)*1] + 1.82 [ln(500)*ln(100)] = +19.57 %

ΔY = 27.41% – 19.57%= **+7.84%**

In this scenario the increase of distance makes the airplane more likely. The reason is that we are now considering a time sensible product. Time has a strong effect in the decisional process because it's an insurmountable constraint: if the delivery requires rapidity and the distance increase, the only available transport mode becomes "air shipping" (even if it's much more costly and the value/weight is low).

Even if Value/Weight is very low, Y increases with distance. From the single interdependency analysis above, we would have expected something different, since we said (analyzing the V/W-distance interdependency) that for V/W behind 391 €/tons Y decrease for higher distances. However we are studying here a "triple interdependency" among the most influent variables: V/W, timeliness and distance and, as shown in table 47, "ln(distance)" has a quite strong and positive coefficient even for products with no time constraints (3,89). The result here is the outcome of a complex synergic effect.

Scenario N.3 (High-0)

Distance "High":

Y= -9.55 * ln(6000) + 4.32 [ln(6000)*0] + 1.82 [ln(6000)*ln(2000)] = +37.26%

Distance "Low":

Y= -9.55 * ln(500) + 4.32 [ln(500)*0] + 1.82 [ln(500)*ln(2000)] = +26%

ΔY = 37.26% – 26%= **+11.26%**

The product is not time sensitive but the probability (Y%) increases thanks to the high value/weight. There are two main logic reasons for this. First of all, in case of high value/weight the impact of transportation cost is lower (there is a lower ad-valorem cost) and airplane becomes more affordable. Moreover, since the higher the distance the higher the transport time, the "sea shipping" becomes less convenient because the value of goods is "chained" in the products for longer time (higher "opportunity costs").

Scenario N.4 (High-1)

Distance "High":

Y= -9.55 * ln(6000) + 4.32 [ln(6000)*1] + 1.82 [ln(6000)*ln(2000)] = +74.85%

Distance "Low":

Y= -9.55 * ln(500) + 4.32 [ln(500)*1] + 1.82 [ln(500)*ln(2000)] = +53,42%

ΔY = 74.85% – 53.42%= **+21.43%**

If distance increases, probability of selecting plane increases by 21%! In this scenario we are considering a valuable good that is also time sensible: the concurrent presence of "need for speed" and high value/weight makes airplane the favorite transportation mode.

In our model we introduced the binary "Time Sensitiveness" variable in order to explicit the role of time in transportation choices. As we have already noted, "Time Sensitiveness" has a strong correlation with "Distance" (and so with "Time Difference"). Moreover, in our model we included another less intuitive correlation between "Time Sensitiveness" and "Value/Weight".

Now we analyze in which way the decisional process about transportation mode is influenced by the presence of a time constraint. In particular, we implement again the scenarios' analysis by developing an analogous matrix as in the case of

"Distance" (we will look now at slacks in the different scenarios, making timeliness vary).

SCENARIOS ANALYSIS 2

	DISTANCE	
VALUE/WEIGHT	Low; Low (1^)	Low; High (2^)
	High; Low (3^)	High; High (4^)

We consider:

- V/W = "Low" = 1000€/ton
- V/W = "High" = 8000 €/ton
- Distance = "Low" = 6565 km (Indian Distance)
- Distance = "High" = 15935 km (Australian Distance)

For each scenario we detect the "Time Sensitiveness" relevance by supposing to change from a no time sensitive product to a time sensitive one:

$\Delta Y = f(\Delta \text{Time Sensitiveness}) = f(\text{Time Sensitiveness} = 1) - f(\text{Time Sensitiveness} = 0)$

$\boldsymbol{\Delta Y} = [-104.73 * (\text{Time Sensitiveness} = 1) + 4.32\ (\ln(\text{Distance}) * \text{Time Sensitiveness} = 1) + 9.09\ (\ln(V/W) * \text{Time Sensitiveness} = 1)] - [-104.73 * (\text{Time Sensitiveness} = 0) + 4.32\ (\ln(\text{Distance}) * \text{Time Sensitiveness} = 0) + 9.09\ (\ln(W/V) * \text{Time Sensitiveness} = 0)]$
$\boldsymbol{= [-104.73 * 1 + 4.32\ (\ln(\text{Distance}) * 1) + 9.09\ (\ln(W/V) * 1)]}$

Scenario No.1 (Low-Low)

$\Delta Y = [-104.73 * 1 + 4.32\ (\ln(6565) * 1) + 9.09\ (\ln(1000) * 1)] = -4.14\ \%$

Scenario No.2 (Low-High)

$\Delta Y = [-104.73 * 1 + 4.32\ (\ln(15935) * 1) + 9.09\ (\ln(1000) * 1)] = -0.33\ \%$

Scenario No.3 (High-Low)

ΔY= [-104.73 1 + 4.32 (ln(6565)* 1) + 9.09 (ln(8000) * 1)]=14.75%*

Scenario No.4 (High-High)

ΔY= [-104.73 1 + 4.32 (ln(15935)* 1) + 9.09 (ln(8000) * 1)]=18.57%*

The first two scenarios give a negative differential of probability. The reason is imputable to the low value/weight that makes the "air shipping" costs unsustainable. However, the "probability difference" is less negative in case of higher distances (scenario N.2): the reason is that in case of longer distance the "need for speeds" becomes more valuable and air transportation becomes more likely.

In the last two scenarios the differential of probability becomes positive. This evidence confirms the strong "role of time" in case of time constraints. As in scenarios N.2 also for scenarios N.4 the probability is higher due to the higher relevance of time for longer distances.

In conclusion, we can note the strong role played by the "Value/Weight" through this scenario analysis. However, results are heavily affected by the multiple interdependencies among the three variables: we analyzed pairs of interdependencies above, finding critical points but these results are not true in absolute values, since they represent an average obtained without making explicit in the equation other explanatory effects.

Even if critical points can be significant to understand tendencies, we need to further investigate on the role of multiple interdependencies (among the three or the four main variables) in next chapters.

5.2.5 STEP (4): Model Improvements

With the final multiple regressions we have reached a good fitting level. However, it's possible to improve further our model.

INTRODUCTION OF "FARE DIFFERENCE 3" AND ITS INTERDEPENDENCIES

First of all, "Fares Difference" indicator can be largely improved by substituting "Fare Difference 2" with the new "Fare Difference 3". This new indicator is much more structured than the previous ones because it takes into account the actual quantity of each single observation (it's not the same to transport 50kg or 50 tons of potatoes) and it's based on real data (Artsana estimation for sea fares and UPS fares for **air cargo** shipping in B2B). We describe in details this variable in chapter 4.2.3 "Independent Variables".

We find that the most fitting function for "Fare Difference 3" is the logarithmic one: for both the linear (regression (1), table 60) and the quadratic function (regression (2), table 60), the coefficients are not significant, while the logarithmic one is consistent and significant. We said above we would expect a negative sign: the higher the fare the lower the likelihood. As for "Fare Difference 2", this new indicator is adept at capturing the products' density effect, more than the distance one (we already discovered how distance doesn't seem to matter from a fare difference perspective because "Fares Difference 1" results always statistically irrelevant).

REGRESSOR		(1)	(2)	(3)
Fares 3	Coef.	-8.18e-12	-3.72e-11	
	SE	6.13e-12	1.73e-11	
	t	-1.33	-2.15	
Sq fares3	Coef.		1.31e-23	
	SE		7.31e-24	
	t		1.79	
Ln Fares3	Coef.			-1.766861
	SE			.1167847
	t			-15.13
R-Squared Adjusted		**0.0003**	**0.0013**	**0.0916**
Root MSE		20.728	20.718	18.778

Table 60 – "Fares Difference (3)" Non-Linear Functions

We compute the correlation coefficient between "Distance" and "Fares Difference 3" in order to get further evidence about their very low relationship. As we expected, the correlation coefficient is actually very low: -0.0414. The basic concept is that distance is not the main driver for fare setting but there are many other costs voices that influence heavily the final fare. For example, the infrastructure level of a harbor (or airport) could strongly influence the fare; if the harbor is old and it doesn't have facilities to support the containers management, time required for loading/unloading operations will be higher than for a modern harbor. As a consequence, there will be also higher costs. The frequency of trade and the "economic relevance" of a certain destination is for sure another relevant cost determinant; it's much more costly to deliver a certain good towards a destination that is not very served because the logistic companies will propose a flight with a low saturation or, at worst they have to arrange a new unscheduled and more costly flight.

Since "Fare Difference 3" includes both the "density" and the "bulky" effects, we exclude them from the model in order to avoid "co-linearity" problems. In particular the variable has been developed by considering both the weight and the volume of each single observation. Moreover, the "saturation level" that we introduced in the variable computation captures the limit of volume for bulky goods.

As we made in the previous cases, we test the most likely interdependences between "Fares Difference 3" and other variables.

REGRESSOR		(1)	(2)	(3)
ln Value/Weight	Coef.		14.4837	
	SE		1.021995	
	t		14.17	
Ln Distance	Coef.			19.49036
	SE			2.32903
	t			8.37
Ln Fares3	Coef.	-1.543742	2.78722	5.629421
	SE	.1283059	.4214325	1.081665
	t	-12.03	6.61	5.20
Time Sensitiveness	Coef.	53.55626		
	SE	8.557226		
	t	6.26		
Int Fares3 - Timeliness	Coef.	-2.116566		
	SE	.4775741		
	t	-4.43		
Int Fares3- Value/weight	Coef.		-.5347053	
	SE		.0565052	
	t		-9.46	
Int Fares3 - Distance	Coef.			-.8608649
	SE			.1267528
	t			-6.79
R-Squared Adjusted		**0.1454**	**0.2806**	**0.1421**
Root MSE		18.981	16.71	18.249

Table 61 - Simple Regressions including Inderdependencies with "Fare Difference (3)"

In regression (1) of table 61, we test the interdependence with "Time Sensitiveness" and we discover that it's significant. The coefficient seems to be unintuitive because the interdependence dampens the probability of selecting "air transport" in case of time sensitive product (the coefficient is negative). However, a deeper derivatives analysis can help us to interpret better the coefficient.

$$Y = 35.33 + 53.5\,X_1 - 1.54\,X_2 - 2.1\,[X_1 * X_2]$$

Where:

X_1 = "Timeliness"

X_2 = "ln(Fares3)"

$$\delta Y/\delta X_1 = 53.5 - 2.1\,X_2$$

$Y' = 0 \rightarrow X_2 = 53.5/2.1 = 25.45 \rightarrow \ln(Fares3) = 25.45 \rightarrow e^{25.45} =$ **115.206.526.520** €

The fare boundary that we find out (115.206.526.520 €) is an extreme value and there are just few observations with a higher or equal fare difference. So, we can coherently affirm that the probability of taking airplanes always increases if we add a time constraint.

In regression (2) of table 61, we test the interdependency between "Fares3" and "Value/Weight". Also in this case the interaction is strongly consistent but the sign is counterintuitive: considering a more valuable good the interdependency's coefficient attenuates the probability (the coefficient is negative). As before, we deeper investigate by going through a derivative's analysis:

$Y = -77.26 + 14.4837\, X_1 + 2.78\, X_2 - 0.53[X_1 * X_2]$

Where: X_1 = "ln(Value/Weight)"
X_2 = "ln(Fares3)"

$\delta Y/\delta X_2 = 2.78 - 0.53\, X_1$

$Y' = 0 \rightarrow X_1 = 2.78/0.5347 = 5.1991 \rightarrow \ln(V/W) = 25.45 \rightarrow e^{5.1991} =$ **181** €/ton

For very low value weights (<181 €/ton) we would have positive derivate. It means that for this kind of "cheap" products, Y increases for higher fares. It sounds weird and we check for the value of Y substituting the threshold 181 €/ton.

$Y = -77.26 + 14*5.1999 = -5\%$

We find a negative probability! Actually cheap products (there are only a few observations under 181 €/ton) are never transported by air and the negative Y is result of approximation in the regressions. For all of those products with value/weight higher than 181 €/ton the probability is negative as expected. So,

products with any “price” higher than 181 €/ton register an airplanes usage that **decreases** for higher fares difference.

The slope is **steeper** for higher value/weights: this sounds somehow strange, since we expect more expensive products to be less sensitive to transportation costs. However, slope by itself is not a measure of airplanes likelihood. We substitute now some different value/weight values, to check for the different fares sensitiveness (table 62).

Value/Weight [€/ton]	Slope	Y% (Fare A= 4112634€)	Y% (Fare B= 81763760€)	Y% (Fare C= 808692820€)
529	-0.54	5.25%	3.6%	2.4%
1462	-1.08	11.78%	8.5%	6%
3981	-1.61	18.2%	13.37%	9.68%

Table 62 - "Air Shipping" probability depending on Value/Weight

The selected values represent the first/second/third quartile for “Value/Weight” and “Fare Difference 3” distributions (see table 63).

	First quartile	Second Quartile	Third Quartile
Value/Weight	529 [€/ton]	1462 [€/ton]	3981 [€/ton]
Fares Difference	4112634€	81763760€	808692820€

Table 63 – Quartiles

Absolute slope increase for higher value-weights, but this is correct because probability is much more sensible to value/weight differences than to fares differences. So, considering for instance “Fare Difference 3”= 808692820 € (that is a very high value corresponding to the third quartile), probability is still higher for higher value/weights even if slope is more negative. We go through a last test to check for consistency of what we found. We compute the value/weight that makes Y=6.1% (how do things change passing from 6% to 6,1%?).

6.1%=-77.26 + 2.78*ln(FareC) + 14.48*lnX – 0.53*lnX*ln(FareC)
X= 1465 €/ton

The more negative slope doesn't imply lower value/weights. So results are coherent and the slope trend results negative because of the logarithmic functions and the very high sensitiveness of Y to Value/Weight.

In regression (3) of table 61, we test the interdependency between "Fares Difference 3" and "Distance". As we expected, there is a strong interdependency.

Now, we test the complete multiple regression by introducing "Fares Difference 3" and all of the interactions.

REGRESSOR		(1)	(2)	(3)
ln Value/Weight	Coef.	9.882791	-3.385465	-1.061227
	SE	1.101697	2.128045	2.186157
	t	8.97	-1.59	-0.49
Ln Distance	Coef.		-10.0114	1.686194
	SE		1.667834	3.181366
	t		-6.00	0.53
Time Difference	Coef.	-.7293365		
	SE	.1490271		
	t	-4.89		
Ln Fares3	Coef.	2.031507	2.169934	6.589717
	SE	.427822	.4186402	1.106798
	t	4.75	5.18	5.95
Airport Density	Coef.	.279629	.2601488	.2753576
	SE	.1144958	.1099047	.1094749
	t	2.44	2.37	2.52
Sq Airport density	Coef.	-.0078544	-.005162	-.0055855
	SE	.0024253	.0024924	.0024833
	t	-3.24	-2.07	-2.25
Time Sensitiveness	Coef.	-52.43985	-72.70082	-76.89549
	SE	11.81534	16.66123	16.61597
	t	-4.44	-4.36	-4.63
Int Fare 3- Value/weight	Coef.	-.4203429	-.4369811	-.4198584
	SE	.0581836	.0569723	.0568589
	t	-7.22	-7.67	-7.38
Int Time Difference-Value/weight	Coef.	.1167914		
	SE	.0199468		
	t	5.86		
Int Distance – Value Weight			1.802249	1.492448
			.2193919	.2299405
			8.21	6.49
Int Fares3 - Timeliness	Coef.	-.7859827	-.6065801	-.5667658
	SE	.4331078	.4255051	.4237215
	t	-1.81	-1.43	-1.34
Int Distance-Timeliness	Coef.		2.567563	2.969726
	SE		1.433901	1.430596
	t		1.79	2.08
Int Time Difference-Timeliness	Coef.	.008748		
	SE	.1170631		
	t	0.07		
Int Value/Weight - Timeliness	Coef.	8.678878	8.139019	8.134537
	SE	.8986141	.8860051	.8820823
	t	9.66	9.19	9.22
Int Fare3 - Distance	Coef.			-.5355674
	SE			.1242437
	t			-4.31
R-Squared Adjusted		**0.3624**	**0.3814**	**0.3869**
Root MSE		16.395	16.148	16.076

Table 64 - Multiple Regressions with all of the"Fares Difference 3" Interdependencies

The first two regressions are the same but we consider “Time Difference” in the first one and “Distance” in the second one. Also in this case, adjusted R^2 is higher with “Distance” than with “Time Difference”. **The reason is that**

"Distance" doesn't include just the time effect but it is able to capture other explanatory factors like "need for coordination", "information exchange", "uncertainty" and "regionalism". From the traditional economic literature about the puzzling effect of distance, it's clear how it's difficult to make explicit these determinants and how they can be "absorbed" by "Distance" (in regression (3) we add the interaction between "Fares Difference 3" and "Distance").

The next step aiming at improving the regression model is to try to capture other explanatory factors that can solve the "omitted variable" problem.

INTRODUCTION OF "FRAGILITY EFFECT", "COUNTRY EFFECT" AND "IMPORT/EXPORT EFFECT"

First of all, we decided to add another "product variable" that would be able to express both the level of fragility and danger for the specific product category. We develop a new dummy variable that we labeled "Fragility&Dangerous": we assign "1" when the product category could have high level of fragility and/or of dangerousness (e.g. toxicity, pollution risk, flammability, etc…). In Appendix C we report all the assigned values for each product category.

Regression (1) of table 65 includes the new dummy variable: adjusted R^2 improves, the dummy's coefficient is significant (t=6.47) and the sign is coherent. The presence of fragility or/and dangerousness' constraints increases the probability of "air shipping": this result is an evidence of the higher security using "air shipping" rather than other transportation means.

We can note that some single coefficients in regression (1) of table 65 like "lnDistance" and "lnValue/Weight" lose their significance. In order to avoid some misleading consideration, we want to point out that this doesn't imply a

null effect of “Distance” and “Value/Weight” because the effect of these variables is “captured” in all their interdependency coefficients.

In regression (2) we tried to adjust the “Airport Density” variable because it is always not very significant in the final multiple regressions. We corrected by considering the ratio between the number of airports and the total population (instead of the country surface). However, results look almost the same.

Then we decided to include in the regression a further dummy variable in order to make distinction between European “export” and “import”. We assign “1” in case of “export”. This variable can capture some eventual differences between import and export. For example, export transportation fare could differ from import fare considering a same route; that’s the case of trade between China and Europe: from Artsana interview, we discovered that the current Ocean Freight for exporting from the European Union to China is much lower than the import one. The reason is that China needs empty container in order to sustain its huge amount of exportations.

The new dummy variable should be able to capture this and many other import/export peculiarities. From Regression (3) to Regression (13) of tables 65/66, we include the Import/Export variable. The variable is never significant: t value is always lower than 1, and we always accept the null hypothesis H_0. Although the variable seems to be irrelevant in trade pattern, we must note that we are considering aggregated data that cannot show temporary differences between import and export (since we have annual data, monthly fluctuations are not visible). Moreover, we are considering many countries together, while there can be some “import/export” effects related to single countries. We might not to see this effect in the analysis because of the presence of other countries that counterbalance this effect.

Finally, we introduced "country variables" in order to capture all the other possible country features that we neglected so far. Each country has particular characteristics related to its demography, geography and economy. However, it's quite unrealistic to define specific variables for each country peculiarity. So, we create a single dummy variable that summarizes all of these country features.

We implement a regression for each country (from regression (2) to regression (12)) by considering its own country dummy variable; for example, in regression (2) we express the country characteristics of Australia by introducing "Country Dummy" that is "1" for Australian observations, "0" for all the other countries.

The coefficient is quite significant for Australia, India and Morocco: this means that for these countries there are some other neglected characteristics that can impact trade pattern in some way.

Australia is for sure a particular country due to its huge remoteness from the rest of the world: the lack of neighboring countries could modify the traditional economic equilibriums and trade patterns.

Morocco is a developing country: its level of infrastructures and its bureaucratic system are still quite backward and this is a real constraint for international trade. Moreover, other demographic elements such as low literacy rate (56%[30] of people ages 15 and above), limited urban development could have a strong impact over trade pattern.

India is classified as a "lower middle income"[31] country. Although it is one of the fastest growing economies (it's one of the five BRICS nations), India has to cope with poverty, illiteracy, corruption and inadequate public health. All these

[30] World Bank Data (2009)

[31] World Bank Data (2010)

demographic factors can strongly influence international trade. Besides, India has comparative advantage in labour-intensive goods and it's largely specialized in clothing production (from preliminary analysis about European trade, we discovered that the 26% of the total European import from India is clothing); since clothing sector could be very time sensitive, "air shipping" could results more used in India than in other similar countries that are not specialized in clothing.

REGRESSOR		(1)	(2)	(3) AUSTRALIA	(4) BRAZIL	(5) CANADA	(6) CHINA	(7) INDIA
ln Value/Weight	Coef.	-.9734366	-1.380165	-1.278772	-.9654314	-.9974669	-1.018004	-1.065389
	SE	2.163516	2.182628	2.164273	2.167492	2.163908	2.164216	2.157355
	t	-0.45	-0.63	-0.59	-0.45	-0.46	-0.47	-0.49
Ln Distance	Coef.	2.025566	1.728119	3.474125	2.02075	1.937616	2.013948	1.740269
	SE	3.147514	3.152003	3.200797	3.157307	3.15488	3.154369	3.14552
	t	0.64	0.55	1.09	0.64	0.61	0.64	0.55
Ln Fares3	Coef.	6.732339	6.761329	6.935146	6.723261	6.758666	6.711986	6.822531
	SE	1.09798	1.109632	1.102777	1.103526	1.101544	1.101362	1.098193
	t	6.13	6.09	6.29	6.09	6.14	6.09	6.21
Airport Density	Coef.	.1783686		.073505	.1768185	.2022411	.130354	.2495924
	SE	.1200616		.1266578	.1219595	.1216917	.1267199	.1213063
	t	1.49		0.58	1.45	1.66	1.03	2.06
Sq Airport density	Coef.	-.0035627		-.0006026	-.0035242	-.0040185	-.0025491	-.0049326
	SE	.0026715		.0029055	.0027253	.0026986	.0028053	.0026905
	t	-1.33		-0.21	-1.29	-1.49	-0.91	-1.83
Time Sensitiveness	Coef.	-77.10479	-75.86296	-77.63573	-77.10955	-77.40398	-76.99155	-77.44736
	SE	16.43941	16.60593	16.42151	16.44837	16.44336	16.44226	16.39159
	t	-4.69	-4.57	-4.73	-4.69	-4.71	-4.68	-4.72
GDP	Coef.	1.86e-07	2.59e-07	9.69e-08	1.85e-07	2.05e-07	2.52e-07	1.70e-07
	SE	9.80e-08	9.08e-08	1.04e-07	9.96e-08	9.93e-08	1.13e-07	9.79e-08
	t	1.90	2.85	0.93	1.86	2.07	2.24	1.74
Fragility&Dangerous	Coef.	7.63148		7.63785	7.632284	7.633155	7.628268	7.650392
	SE	1.178999		1.177635	1.179631	1.179162	1.1792	1.175572
	t	6.47		6.49	6.47	6.47	6.47	6.51
Int Fare 3-Value/weight	Coef.	-.4211244	-.4161034	-.4214157	-.4213473	-.4219293	-.4205491	-.4278295
	SE	.0562459	.0567875	.0562086	.0563315	.0562844	.0562861	.0561382
	t	-7.49	-7.33	-7.50	-7.48	-7.50	-7.47	-7.62
Int Distance – Value Weight	Coef.	1.47671	1.525215	1.515413	1.476121	1.480076	1.481457	1.504505
	SE	.2275678	.2296608	.2278048	.2279669	.2276225	.2276474	.2270365
	t	6.49	6.64	6.65	6.48	6.50	6.51	6.63
Int Fares3 – Timeliness	Coef.	-.4057228	-.6074177	-.359779	-.4062679	-.3905982	-.4036348	-.3756135
	SE	.4202153	.4233202	.4201056	.4204856	.4204515	.4202864	.4190669
	t	-0.97	-1.43	-0.86	-0.97	-0.93	-0.96	-0.90
Int Distance-Timeliness	Coef.	3.244003	2.896489	3.204462	3.244265	3.245603	3.240981	3.215952
	SE	1.415933	1.42975	1.41439	1.416758	1.416147	1.416189	1.411842
	t	2.29	2.03	2.27	2.29	2.29	2.29	2.28
Int Value/Weight – Timeliness	Coef.	7.570703	8.164939	7.571522	7.572192	7.573901	7.554961	7.575194
	SE	.8769286	.8817254	.8759086	.8775849	.877049	.8771755	.8743687
	t	8.63	9.26	8.64	8.63	8.64	8.61	8.66
Int Fare3 - Distance	Coef.	-.5609096	-.5642216	-.5857532	-.5596468	-.5627582	-.5593646	-.5652955
	SE	.1234865	.1247539	.1241499	.1243805	.123933	.1239233	.1235462
	t	-4.54	-4.52	-4.72	-4.50	-4.54	-4.51	-4.58
Airport Density Adjusted	Coef.		-.2511018					
	SE		.3373316					
	t		-0.74					
Import/Export Dummy	Coef.			-.0360768	-.0422339	-.0368443	-.0436779	-.0276581
	SE			.7205486	.7217726	.7214945	.7215016	.7192882
	t			-0.05	-0.06	-0.05	-0.06	-0.04
Country Dummy	**Coef.**			**-4.498733**	**-.0917705**	**1.621887**	**-1.690955**	**4.764897**
	SE			**1.746492**	**1.402288**	**1.315437**	**1.426896**	**1.292246**
	t			**-2.58**	**-0.07**	**1.23**	**-1.19**	**3.69**
R-Squared Adjusted		**0.4001**	**0.3874**	**0.4015**	**0.3995**	**0.3999**	**0.3999**	**0.4036**
Root MSE		15.903	16.07	15.884	15.911	15.905	15.905	15.856

Table 65 - Complete Multiple Regressions - With Country Dummy (a)

REGRESSOR		(8) JAPAN	(9) MOROCCO	(10) RUSSIA	(11) SWITZERLAND	(12) TURKEY	(13) USA
ln Value/Weight	Coef.	-.9906508	-1.000925	-.9672056	-.9906484	-.988984	-.9716787
	SE	2.164603	2.163855	2.164624	2.164689	2.164877	2.164554
	t	-0.46	-0.46	-0.45	-0.46	-0.46	-0.45
Ln Distance	Coef.	1.717084	1.578194	1.972414	1.727037	1.829189	1.967065
	SE	3.188671	3.172337	3.157324	3.193937	3.184154	3.157287
	t	0.54	0.50	0.62	0.54	0.57	0.62
Ln Fares3	Coef.	6.709146	6.687766	6.756992	6.704935	6.727037	6.7313
	SE	1.101929	1.101649	1.104479	1.102271	1.101626	1.101666
	t	6.09	6.07	6.12	6.08	6.11	6.11
Airport Density	Coef.	.0160504	.1985546	.1685505	-.0244222	.2025845	.1841023
	SE	.2791594	.1211553	.1228986	.3703726	.1327806	.1209728
	t	0.06	1.64	1.37	-0.07	1.53	1.52
Sq Airport density	Coef.	-.0005184	-.0044138	-.0034829	.005938	-.0042689	-.0036845
	SE	.005429	.0027542	.0026813	.0166379	.0031389	.0026901
	t	-0.10	-1.60	-1.30	0.36	-1.36	-1.37
Time Sensitiveness	Coef.	-77.2458	-77.30356	-77.06598	-77.2236	-77.13268	-77.17815
	SE	16.44768	16.44173	16.44756	16.44784	16.44726	16.44814
	t	-4.70	-4.70	-4.69	-4.70	-4.69	-4.69
GDP	Coef.	2.55e-07	1.59e-07	1.90e-07	2.42e-07	1.74e-07	2.32e-07
	SE	1.45e-07	1.00e-07	9.87e-08	1.37e-07	1.02e-07	1.47e-07
	t	1.76	1.58	1.93	1.76	1.70	1.58
Fragility&Dangerous	Coef.	7.636254	7.623614	7.631715	7.635478	7.635532	7.633808
	SE	1.179512	1.17914	1.179575	1.179533	1.179591	1.179572
	t	6.47	6.47	6.47	6.47	6.47	6.47
Int Fare 3-Value/weight	Coef.	-.4213944	-.4216336	-.4214859	-.4212459	-.4211041	-.4216257
	SE	.0562979	.0562805	.0563054	.0562985	.0563014	.0563085
	t	-7.49	-7.49	-7.49	-7.48	-7.48	-7.49
Int Distance – Value Weight	Coef.	1.478785	1.48072	1.477048	1.478489	1.478378	1.477399
	SE	.2276913	.2276199	.227688	.2276937	.2277122	.2276888
	t	6.49	6.51	6.49	6.49	6.49	6.49
Int Fares3 – Timeliness	Coef.	-.4049877	-.3955739	-.4041111	-.4054733	-.4062296	-.4046744
	SE	.4203901	.4203318	.4204426	.4203971	.4204134	.4204222
	t	-0.96	-0.94	-0.96	-0.96	-0.97	-0.96
Int Distance-Timeliness	Coef.	3.238298	3.223446	3.2465	3.239002	3.24429	3.240494
	SE	1.416566	1.416184	1.416668	1.416593	1.416628	1.416645
	t	2.29	2.28	2.29	2.29	2.29	2.29
Int Value/Weight – Timeliness	Coef.	7.591302	7.594538	7.560028	7.588988	7.574136	7.580665
	SE	.8778616	.8772086	.8778423	.8778635	.8773738	.8776456
	t	8.65	8.66	8.61	8.64	8.63	8.64
Int Fare3 - Distance	Coef.	-.5581022	-.5557219	-.563414	-.557747	-.5603708	-.5604145
	SE	.1239995	.1239645	.1242388	.1240337	.1239592	.1239596
	t	-4.50	-4.48	-4.53	-4.50	-4.52	-4.52
Import/Export Dummy	Coef.	-.0451492	-.0717648	-.0368813	-.0453555	-.0374184	-.0440453
	SE	.721694	.7218144	.7218919	.7217136	.7218216	.7217346
	t	-0.06	-0.10	-0.05	-0.06	-0.05	-0.06
Country Dummy	**Coef.**	**2.09295**	**-1.945827**	**-.5175157**	**-13.22016**	**-.6915676**	**-.8499729**
	SE	**3.252771**	**1.514513**	**1.386252**	**22.86003**	**1.601418**	**2.012562**
	t	**0.64**	**-1.28**	**-0.37**	**-0.58**	**-0.43**	**-0.42**
R-Squared Adjusted		**0.3996**	**0.4000**	**0.3995**	**0.3996**	**0.3995**	**0.3995**
Root MSE		15.909	15.904	15.91	15.909	15.91	15.91

Table 66 - Complete Multiple Regressions - With Country Dummy (b)

5.2.6 Multiple Interdependencies Analysis

We have previously checked the single interdependencies between the most relevant determinants: “Time constraint”, “Fares difference”, “FOB prices” (or “Value/Weight”), “Distance” finding that concurrent effects are always coherent with what expected.

Summarizing:

- FOB price is actually measured as the product categories’ value/weight (look at the variables description chapter for further explanations);
- “Fares Difference 3” is an estimation of the fares really applied by logistic companies weighted on the “trade volume”;
- “Timeliness” is a dummy variable aimed at understanding how time constraints can change the analysis;
- “Distance” is the average distance from EU27 to the given trade partner: we measured distances to Switzerland and Russia as “DOT” (distance weighted to trade) to be more precise, considering that proximity of these countries can undermine results.

From literature, the so called “*Distance Puzzle*” arises: even if transportation costs went down, distance still matters for bilateral trade in gravity models. We said that one of the possible reasons for distance to be still so important for trade is the increasing need for timely delivery. Airplanes can overcome the time limit of trading with far partners, while having very high extra-costs.

We analyzed the “Costs-Value/Weight” trade-off **having the confirmation that likelihood of using planes is much more affected by products prices than by extra-costs** (air shipping probability varies more with value/weight than with fare changes). In the table below (table 67) we propose again the quartiles analysis we made:

Value/Weight [€/ton]	Slope	Y% (Fare A= 4112634 €)	Y% (Fare B= 81763760 €)	Y% (Fare C= 808692820 €)
529	-0.54	5.25%	3.6%	2.4%
1462	-1.08	11.78%	8.5%	6%
3981	-1.61	18.2%	13.37%	9.68%

Table 67 - Y% according to V/W and "Fare Difference (3)"

Indeed, Y is about 10% for products with third quartile value/weight and fares (yellow cell in the table), while there's a similar 11% for products whose value/weight corresponded to the median and fares corresponded to the first quartile (orange cell in the table)!

The point is that for very low extra-costs (first quartile) we would expect Y to "jump up". Actually Y is much higher looking at the same line, but it falls again to 11% considering the median value weight, this confirming the stronger price importance.

We will assess role of **time constraints** and effect of **distance** in framework. For now, we will try to assess the integrated effect of transportation costs and products value while distinguishing for products that have particular time constraints. Then, we will assess the impact of distance. Finally, we will check the results by substituting "Distance" with "Time Difference".

Framework with "timeliness constraint"

From STATA, we report equation linking "Value/Weight", "Fares Difference 3" and "Timeliness" to Y:

$$Y = -58.52 - 53.88 * Timeliness + 11.72 * \ln(V/W) + 2.12 * \ln(Fare\ Difference\ 3) - 0.43 * [\ln(V/W) * \ln(Fare\ Difference\ 3)] - 0.96 * [\ln(Fare\ Difference\ 3) * Timeliness)] + 9.22 * [\ln(V/W) * Timeliness]$$

We put:

X = *ln(V/W))*

Z = *ln(FareDifference3)*

t = *Timeliness*

Y = - 58.52 - 53.88*t + 11.72 *X + 2.12* Z - 0.43*[X*Z] - 0.96*[Z*t] + 9.22*[X*t]

DERIVATES ANALYSIS

As in the previous cases, we make the derivate analysis in order to point out the "critical values". Since there are more than two variables we proceed with **partial derivates**:

$\delta Y/\delta X = 11.72 - 0.43*Z + 9.22*t$

$\delta Y/\delta X \geq 0 \rightarrow Z \geq 27.26 + 21.44*t$

$\delta Y/\delta Z = 2.12 - 0.43*X - 0.96*t$

$\delta Y/\delta Z \geq 0 \rightarrow X \leq 4.93 - 2.23*t$

Considering that "t" is a binary variable [0;1], we distinguish the two scenarios.

t = 0 → "normal product", without restrictive time constraint (evaluated on the basis of product life-cycle, perishability, need for production synchronization).

$\delta Y/\delta X \geq 0 \rightarrow Z \leq 27.26 \rightarrow$ "Fares Difference 3" $\leq 6.9 * 10^{11}$ €

$\delta Y/\delta Z \geq 0 \rightarrow X \leq 4.93 \rightarrow$ "Value/Weight" ≤ 138.4 €/tons

Actually, almost all of the observations in our data-base verify the first equation, while there are only a few observations over thousands that verify the second equation. We can conclude then:

$\delta Y/\delta X > 0 \quad \forall Z$

$\delta Y/\delta Z < 0 \quad \forall X$

t = 1 [products with time constraints]

$\delta Y/\delta X \geq 0 \rightarrow Z \leq 48.7 \rightarrow$ "Fares Difference 3" $\leq 1.4 * 10^{21}$ €

$\delta Y/\delta Z \geq 0 \rightarrow Z \leq 2.7 \rightarrow$ "Value/Weight" ≤ 14.88 €/tons

With time constraint, results are reinforced:

$\delta Y/\delta X > 0 \quad \forall Z$

$\delta Y/\delta Z < 0 \quad \forall X$

With or without time constraints, we have monotonic functions (it was already obvious by the definition of linear regression). However, we understood the derivates signs: **Y increases with "Value/Weight" for any "Fares Difference", while Y decreases if "Fare Difference" increases for any "Value/Weight".**

QUARTILE ANALYSIS

Derivates and slopes are higher if products are time-sensitive. In order to deeper investigate on the different factors' impact, we go through a "**quartile analysis**". The basic idea is to subdivide products in different categories according to the three different "product attributes" (V/W, Fares Difference and Time Sensitiveness); for each product attribute, we point out three reference values that correspond to the first, second and third quartile. Evidently, for the binary "Time Sensitiveness" variable this procedure is not required.

In the following table, we report the quartiles values:

	First quartile	Second Quartile	Third Quartile
Ln(Value/Weight)	6.27	7.29	8.29
Ln(Fares Difference)	15.23	18.22	20.51

Table 68 - Quartile Values

In total we get 18 scenarios. In the following table we compute the predicted probability of selecting "air shipping" by imputing the scenarios values in the main regression.

Ln VW	Ln Fare	Time	Y
6.27	15.23	0	6.190397
6.27	15.23	1	-4.501
6.27	18.22	0	4.467858
6.27	18.22	1	-9.09394
6.27	20.51	0	3.148589
6.27	20.51	1	-12.6116
7.29	15.23	0	11.46492
7.29	15.23	1	10.17792
7.29	18.22	0	8.430966
7.29	18.22	1	4.273566
7.29	20.51	0	6.107303
7.29	20.51	1	-0.2485
8.29	15.23	0	16.63602
8.29	15.23	1	24.56902
8.29	18.22	0	12.31637
8.29	18.22	1	17.37897
8.29	20.51	0	9.008003
8.29	20.51	1	11.8722

Table 69 - Scenarios Analysis According to Different Variables Quartiles

The presence of time constraint lowers Y for the first and second quartile of value/weight. Trend changes for third value/weight quartile: for these products,

for any fare applied, it becomes more likely to move goods by air if they have time constraints.

In order to get graphical evidence, we draw here the probability functions: we put value/weight on the abscise axis while we draw a specific curve for each combination between "Fare" and "Timeliness" (in total we get six scenarios): "timely trade" has lower Y until reaching values that range from 7 to 8 in the horizontal axis.

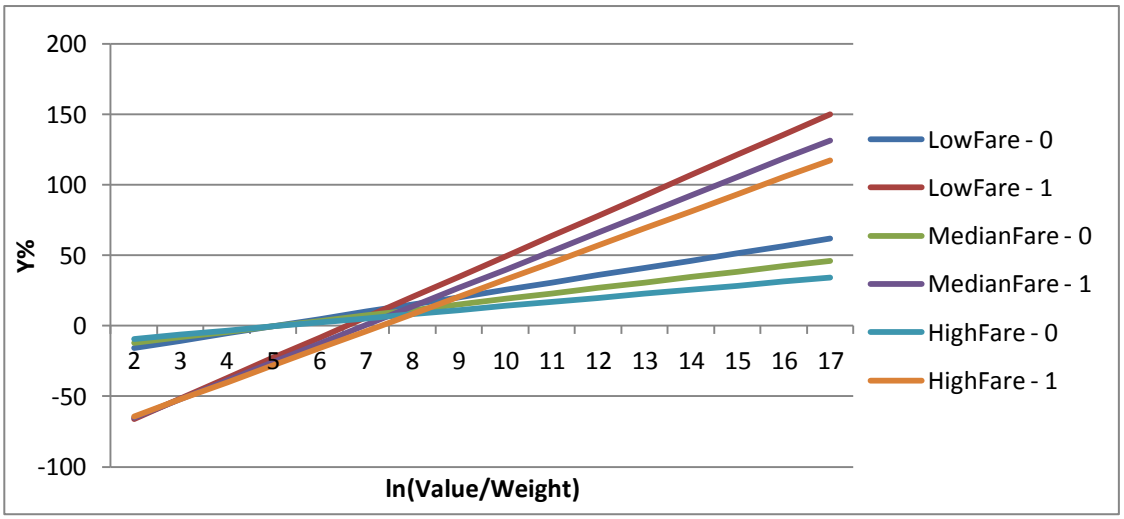


Figure 33 - Y% =f (ln(V/W|Fare3;Timeliness)

There are probabilities smaller than "0" and bigger than "100" in graph (vertical axis). It doesn't make any sense considering that Y represents a probability. However, we remind that these functions are the result of regression analysis: software draws functions in order to minimize R^2. We can consider Y equal to 100, when bigger than "100" and Y equal to "0", when negative.

It could be interesting to define the threshold value/weights for each of six scenarios: these "thresholds" represent those value/weights for which $Y_0 = Y_1$ (in these points "airplanes probability" is the same with or without time constraints).

We compute "the Value/Weight thresholds" by making the intersections between the "time sensitive functions" and the "non-time sensitive functions" of the previous graph (we obtain the intersection points between "LowFare-0" and "LowFare-1", between "MedianFare-0" and "MedianFare-1", between "HighFare-0" and "HighFare-1"). In table below we report the results:

	Ln (Value/Weight)	Value/Weight	Y|t=1	Y|t=0
Ln(Low Fare)	7.429587852	1685.11	12.18674	12.18674
Ln(Medium Fare)	7.740911063	2300.57	10.18294	10.18294
Ln(High Fare)	7.979349241	2920.03	8.106898	8.106898

Table 70 - V/W Thresholds (1)

With "Low Fare" (low transportation costs because of low density), "timely trade" corresponds to a higher Y if value/weight is higher than 1685 €/ton; the threshold grows at 2300 €/ton in relation to the "Median Fare"; it grows at 2920 €/ton considering the "High Fare" (third quartile). The last two yellow columns show the predicted probability in case of "time constraint" (t=1) and "no-time constraint" (t=0): obviously, in correspondence with these "threshold values" the probability is the same.

So, we find that products more expensive than 3000 €/tons are "positively" sensitive to time constraints for any kind of distance or product density (Y grows up if products have time constraints).

Now, we repropose the graph without considering a single "Fare" value (corresponding to a certain quartile) but by taking into account "fares ranges" and average predicted Ys. Fares ranges are subdivided as follow:

A= Low Fare → [15<Ln(FareDifference3)<16];

B= Median Fare → [18<Ln(FareDifference3)<19];

C= High fare → [20.2<Ln(FareDifference3)<20.8];

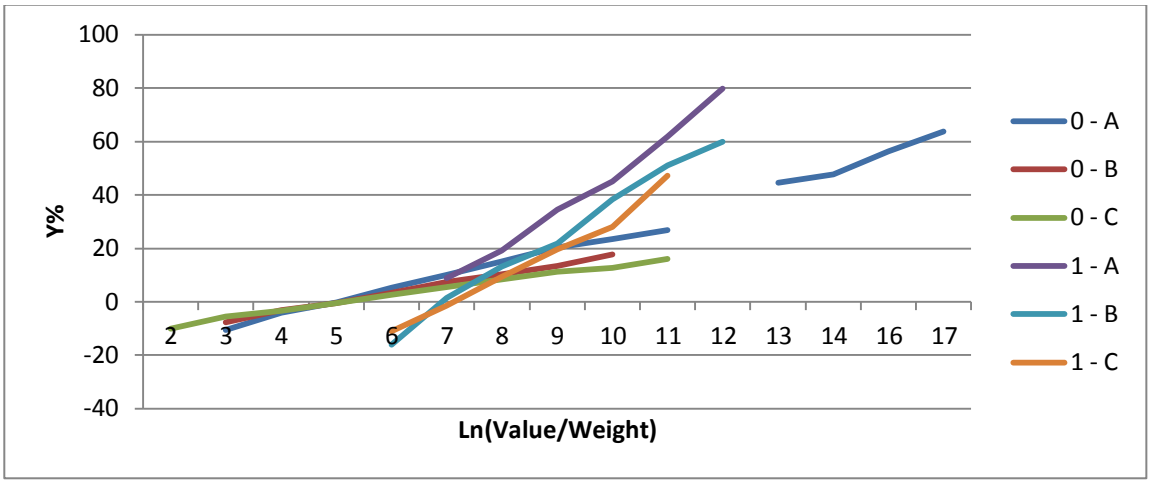


Figure 34 - Y% =f(ln(V/W)|Range of Fare3; Time Sensitiveness)

In the next chart, we replicate the graph including all of the available observations used to build equation: we don't consider "partial ranges" (as in graph 34) but we subdivide all the predicted probabilities in four categories. Here the fares intervals (divided by quartiles values):

Left Tail → [0<Ln(FareDifference3<15.3]

Low Fares → [15.3<Ln(FareDifference3<18.2]

High Fares → [18.2<Ln(FareDifference3<20.51]

Right Tail → [Ln(FareDifference3)>20.51]

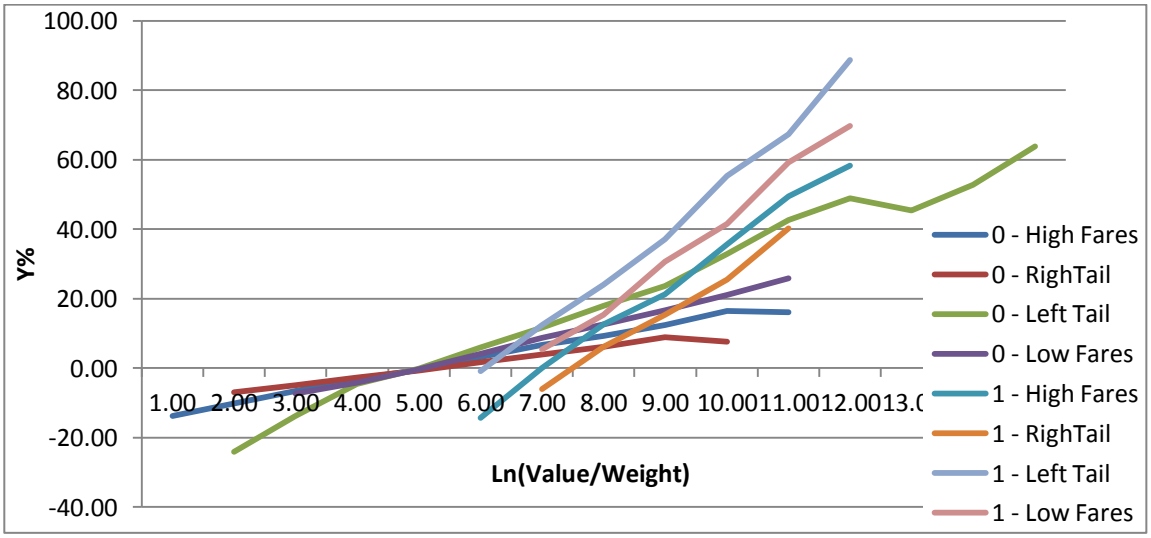


Figure 35 - Y% =f(ln(V/W)|Interval of Fare3; Time Sensitiveness)

Functions are the average of different fares observations in the same category (fares can be very different considering that we express them in logarithms). That's the reason why the lines drawn aren't linear. However, the slope trend is still consistent. That's graphical evidence that Y is more sensitive to V/W than Fare.

From the previous three representations, we get empirical evidences about a same result: *"Value/Weight"* influences relatively more than *"Fare Difference 3"* the likelihood of using the fast mean of transport.

By distinguishing between "standard" and "timely products", we understood this to be "less true" for "normal products". We have confirmation about this from the tables below (table 71 and 72) in which we represent the predicted probabilities distinguishing between t=1 (presence of time constraint) and t=0 (no time constraints). In both the tables, probability differential (ΔY%) increases more (in absolute terms) according to fare than to value/weight; however, the effect is stronger in case of "need for timeliness".

Without time constraint (table 71), if ln(fare) increases from 15.23 (first quartile) to 20.51 (third quartile), the predicted probability varies by |1.52%| (considering first quartile of value/weight); while if ln(value/weight) increases from 6.27 (first quartile) to 8.29 (third quartile), the predicted probability varies by |5.22%| (considering the first quartile of fare difference).

The differential is more marked with the presence of time constraint (table72) according to the previous example, predicted variable variation is |4.05%| in case of fare increase, while it's |14.53%| in case of value/weight rise.

Normal products without particular time constraints:

		Ln(fares)			**ΔY%|ΔLn(Fares)**
		15.23	18.22	20.51	=ΔY%|[ln(High Fare)-ln(Low Fare)]
Ln (V/W)	6.27	**6.19**	**4.47**	**3.14**	-1.52
	7.29	**11.46**	**8.43**	**6.11**	-2.67
	8.29	**16.63**	**12.31**	**9.01**	-3.814
	ΔY%|ΔLn(V/W) =ΔY%|[ln(High V/W)-ln(Low V/W)]	5.22	3.92	2.93	

Table 71 - Y% Slacks

Special products with particular time constraints:

		Ln(fares)			**ΔY%|ΔLn(Fares)**
		15.23	18.22	20.51	=ΔY%|[ln(High Fare)-ln(Low Fare)]
Ln (V/W)	6.27	**-4.50**	**-9.09**	**-12.61**	-4.05
	7.29	**10.18**	**4.27**	**-0.24**	-5.21
	8.29	**24.57**	**17.38**	**11.87**	-6.34
	ΔY%|ΔLn(V/W) =ΔY%|[ln(High V/W)-ln(Low V/W)]	14.53	13.23	12.24	

Table 72 - Y% Slacks

In conclusion, Y is more sensitive to Value/Weight (price) than to fares (costs). This assumption is particularly true for those products with time constraints. These products are never transported by air (less than normal products) if they have low "Value/Weight", while they are more likely to be moved by air for value/weights higher than 3000 €/tons (even despite very high freight costs).

Looking at previous literature, we can assume there is correlation between products value and products level of substitutability: there is always a close trade partner producing cheap products, while it's more difficult to substitute high value products. However, the most

important reason is for sure that freight costs are more easily "sunk" with expensive products.

The combined effect of Distance in Framework

Now we assess the impact of distance in framework. We know that fares computed to define the indicator "FareDifference3" take already into account distance (different distance implies different fares in the given observation). However, we have already demonstrated that **effect of distance** is not captured by freight costs varying with distance. Indeed, the correlation between these two variables is very low (the density component is actually the most relevant for fares definition). Hence, we don't risk co-linearity distortion while considering them concurrently as independent variables: all of the coefficients in equation are statistically consistent.

From STATA, we report equation linking "Value/Weight", "Fares Difference 3", "Timeliness" and "Distance" to Y:

$$\begin{aligned} Y = -72.01 - 1.3 * [\ln (V/W)\ ] + \ 6.48 * [\ln (Fares\ Difference3)\ ] - 76.53 * (Timeliness) + \ 1.63 \\ * [\ln (Distance)\ ] - 0.41 * [\ln (V/W) * \ln (Fare\ Difference3)\ ] - 0.53 \\ * [\ \ln (Fare\ Difference3) * \ln [(Distance)] + 1.5 * [\ln (Distance) * \ln (V/W)\ ] \\ + 2.9 * [\ln (Distance) * Timeliness] + 8.16 * [\ln (V/W) * Timeliness] - 0.58 \\ * [\ \ln (FareDifference3) * Timeliness] \end{aligned}$$

We draw charts reporting regression functions related to "Value/Weight". We develop many graphs according to different distances: we report in the Appendix G all of the graphs we made. We immediately note that distance changes the balance we found before. This means that it has an impact over the probability Y. Now, our aim is to try to figure out in which way "Distance" affects trade patterns.

In order to understand the role of distance we implement a scenarios comparison; firstly, we set two arbitrary values for distance by assuming 1000

km as “Short Distance” and 10000 km as “Long Distance”. Then we draw all of the functions of the different scenarios by putting the ln(Value/Weigh) in the abscises axis and the predicted probability in the ordinates axis. We get in total twelve scenarios: each of them is characterized by a certain distance (1000 or 10000), by the presence or the absence of time constraint (0 or 1) and by a certain Fares Difference (Low, Medium or High). To get more visible results, we divide the scenarios representation in “Non-Timely” scenarios and “Timely” scenarios.

“Non-Timely scenario” (“Normal products” non-time sensitive)

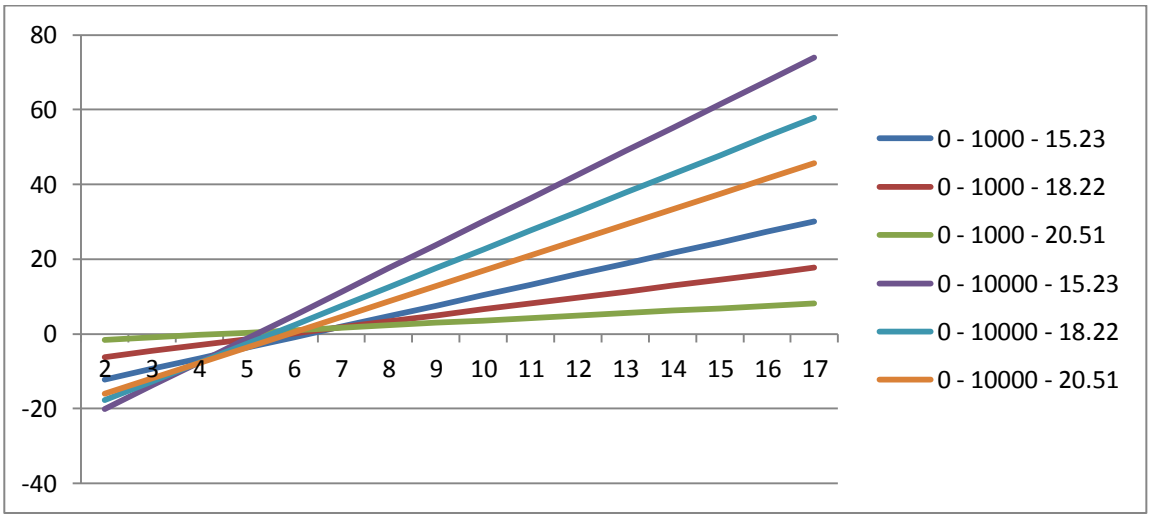


Figure 36 - Y% "Non-Timely Scenario"

For any fare applied, slopes are steeper for the “distant trade” (farer trade has higher Y for almost any value/weight over 400 €/tons).

"Timely Scenario" (Products particularly time-sensitive)

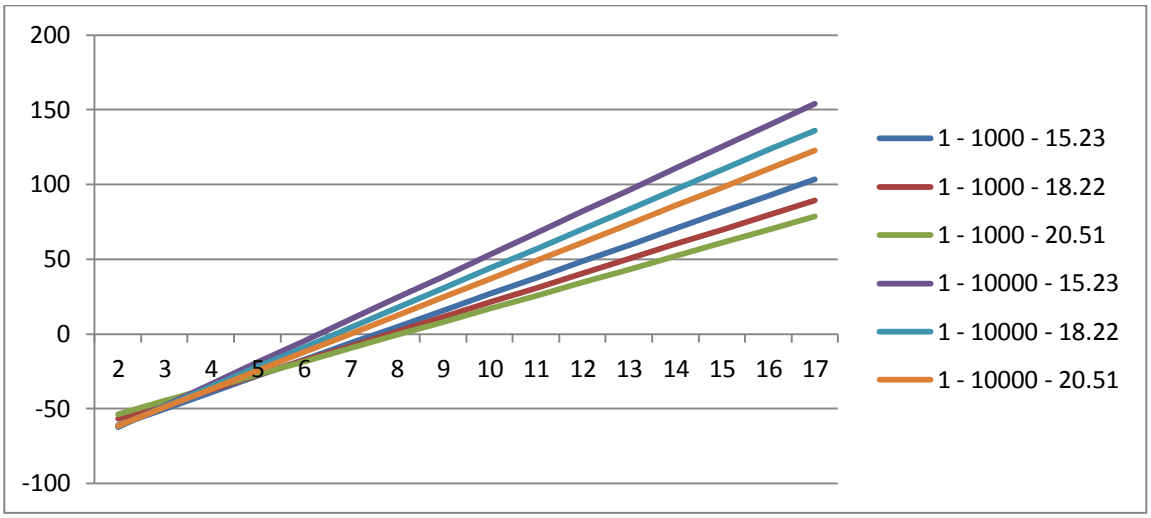


Figure 37 - Y% "Timely Scenario"

Results are consistent in case of time constraints. However, slacks are lower here for both of different fares level and distance.

So, we can say that distance has still an important role in model; now we try to understand which variable has the stronger impact over Y going through the slacks analysis (as already done before). We consider those values corresponding to the quartiles for the different variables.

	Distance	ln(Fare)	ln(Value/Weight)
First Quartile (Low)	2404.77	15.23	6.27
Median	6565.79	18.22	7.29
Third Quartile (High)	9086.04	20.51	8.29

Table 73 - Quartiles Values

Here, the table with the different probabilities, and then the slacks analysis.

	LowFare			MedianFare			HighFare		
	Low VW	Median VW	High VW	Low VW	Median VW	High VW	Low VW	Median VW	High VW
Non-timely									
Low Distance	2.45	6.66	10.80	1.80	4.76	7.67	1.30	3.31	5.28
Median Distance	5.42	11.18	16.82	3.18	7.68	12.10	1.47	5.01	8.49
High Distance	6.38	12.63	18.76	3.63	8.63	13.53	1.52	5.56	9.52
Timely									
Low Distance	0.00	3.36	15.66	0.00	0.00	10.80	0.00	0.00	7.08
Median Distance	0.00	10.79	24.59	0.00	5.56	18.14	0.00	1.56	13.20
High Distance	0.00	13.19	27.48	0.00	7.45	20.51	0.00	3.05	15.18

Table 74 - Y% = f(V/W; Fare3; Time Sensitiveness; Distance)

To understand slacks, we look at change of Y, comparing two variables per time. In the Appendix G, we report all of the tables with slacks analysis. Contributes of Value/Weight and timeliness seem determinant, so that the most important slacks come from the interdependency between these two variables. We already talked about the relationship between Value/Weight and timeliness above: timeliness implies negative ΔY for very low Value/Weight, it implies big positive ΔY for high Value/Weight (the critical point found before was ln(VW)=7,29). This result is even reinforced while considering this new equation with the distance contribute. We said that V/W (price) is more important than exceeding cost in the model, and we verify now that it is more influencing than distance as well.

It can be interesting to define for which products (according to “Value/Weight”, “timeliness”, “density”) they start taking into consideration the airplanes modality (Y>0) making distance of trade vary. The main goal is to determine for which product characteristics “time constraints” reflect higher probabilities of

using airplanes; in particular, we want to assess how these product attributes vary according to distance.

First of all, we can define the "Value/Weight thresholds" looking for the "0" of our functions. They can be considered critical points since functions are monotonic (otherwise a "0" doesn't necessarily imply Y to be positive and grow up linearly for higher X values). It's clear we are dealing with monotonic functions since these equations are the output of OLS regression. Moreover, we can easily have further evidence by looking at all of the graphs we have drawn.

Here the table with **value/weight thresholds** (table 75). We can consider these values as **product price limit**. For value/weights higher than those, "traders" can take into consideration the option of delivering by air, while for lower values, "air shipping" becomes unsustainable. In other words, these values represent the "0" values for our equation (Y=0): it means that even only one euro more in price implies some airplanes usage because air freight costs become bearable.

	Low Fare		Median Fare		High Fare	
Distance	Timely 0	Timely 1	Timely 0	Timely 1	Timely 0	Timely 1
500	1830	3361	1028	4308	7801	5493
1000	558	1948	365	2541	90	3275
2500	287	1091	283	1472	277	1943
5000	219	761	262	1057	332	1426
6000	207	699	259	977	342	1326
7500	195	632	255	892	352	1220
9000	187	585	252	831	360	1143
10000	182	559	251	799	364	1102
12000	176	520	249	748	370	1037
12500	174	511	248	737	371	1024
15000	168	477	246	692	376	967

Table 75 - Value/Weight Tresholds

To have a graphical support we report results in a graph by putting "Distance" in the horizontal axis and "Value/Weight" in the vertical one. We represent six

different scenarios according to all of the possible combinations between "Fares Difference" (Low, Medium and High) and "Time Sensitiveness" ("0" or "1").

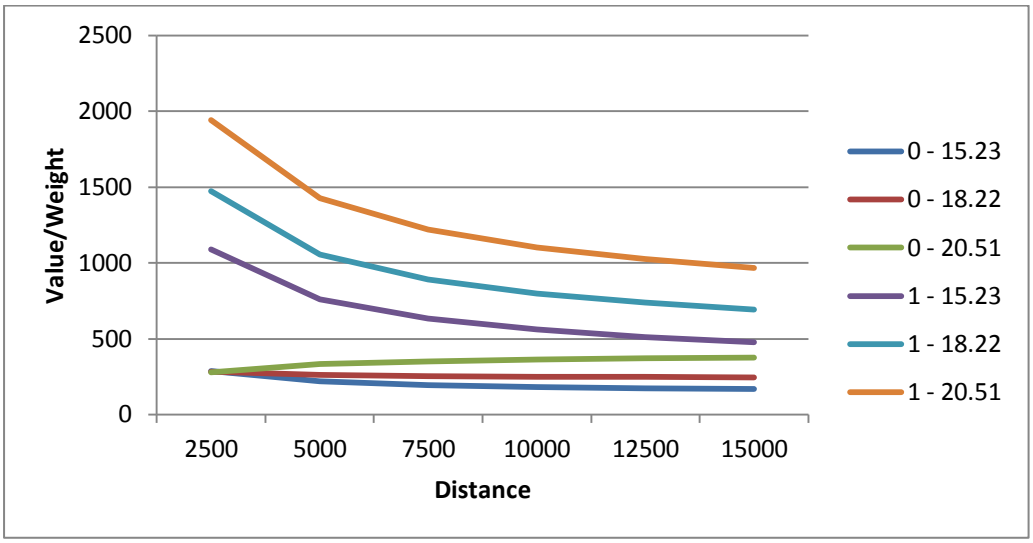


Figure 38 - Threshold Values/Weights = f(Distance)

We discover that minimum price values to use airplanes decrease with distance. A reasonable explanation for this evidence could be the **higher "need for speed" for longer distances**. Indeed, "time advantage" that we got by using airplanes widens over longer distances and airplane results more advantageous for a bigger range of products (even cheaper products). The point is that even if "fares difference" increases a little with distance and from a financial perspective we wouldn't have expected a higher capacity of sinking transportation costs for longer distances, the "**premium price**" (that the final customer is willing to pay for getting more quickly deliveries) makes the airplanes sustainable.

Obviously, the slope of the curves is much steeper in case of need of timeliness: time is not important in absolute values, but it depends on the relative importance that traders give to it. In particular, we see a strong change of slope for all of the "three time sensitive lines" where distance is close to 5000 Km. This distance can represent the closest "trans-oceanic" routes. So time-advantages are enforced by the fact that the alternative mean of transport is the slow ship.

There is only one line that has opposite trend (the green line in chart). It refers to observations with no time constraint and very high "Fares Difference". We already know that "fares slack" is especially influenced by freight volume (if volume and frequency of trade is very high, airplanes become too expensive and "air shipping" is used only for urgencies) and freight density (the higher the density the higher the fares difference). According to this, we can suppose that we are talking about products that are widely traded (high total volume, high trade frequency) and very "heavy", this resulting in strong not sustainable extra-costs. For these products categories, companies can start to take into consideration the idea of delivering by air, only if FOB price is very high (so that cost can be sunk). **The positive trend could be explained considering that for very big volume of trade, distance start being an important cost component**. Evidently, distance matters from a financial perspective only if delivery frequency is extremely high. Indeed, in this case longer distance implies more expensive goods for transportation costs to be sunk.

To conclude, distance has still an impact on transportation costs if volume of trade and delivery frequency are very high. In this scenario the "time advantage" of airplanes increases slower than extra-costs for longer distances. We have to underline that this anomalous line in chart refers to "non-timely trade": once again impact of time doesn't have an absolute value, but it depends on the relative importance for consumers.

Finally, looking at the red line, the slope is almost null. Evidently, for this kind of products, distances extra-costs equal time advantages and airplanes fare can be "sunk" in a same given price level for any distance.

Considering absolute values, "timely lines" are above "non-timely lines" on average. It can be explained by looking at average Value/Weight for the two product categories. Generally, products with time constraints have much higher value/weight. Many "timely products" are composed by different components (that require very short delivery times because of synchronization of the

production process) and, as a consequence they have higher production costs translating in higher final prices. Moreover, we must consider that many goods have time constraints because of their very short "life cycle": final consumers are willing to pay higher prices for timely delivery. So, there is close relationship between timeliness and value/weight.

So far we made the analysis by defining those limit "prices" for which airplanes become sustainable (Y=0). Let's now compare value/weights corresponding to a 0.3 Y (for these products the likelihood of using planes is 30%). As in the previous case, we studies different scenarios by varying the different product attributes. To make the analysis clearer, we report three distinct graphs, one for any different fare level: in each chart we make the comparison between the minimum value/weights (vertical axis) corresponding to a null Y (Y=0) and the minimum value/weights that are associated to 30% Y (Y=30%).

GRAPH A: Low Fare = $e^{15.23}$ (First Quartile)

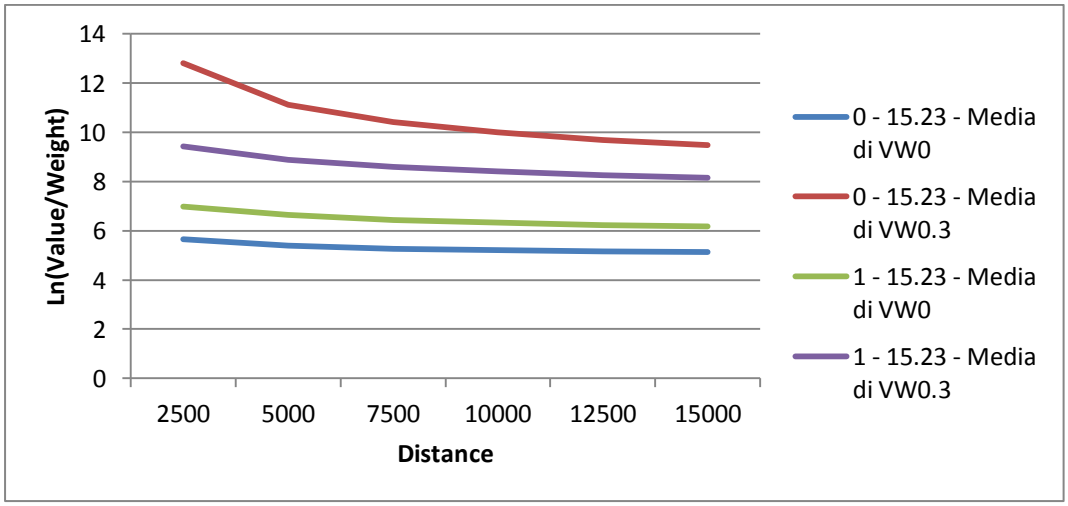


Figure 39 - ln(Thresholds V/W)=f(Distance) --> for Y=0 and Y=0.3 (Low Fare3)

GRAPH B: Median Fare = $e^{18.22}$ (Median)

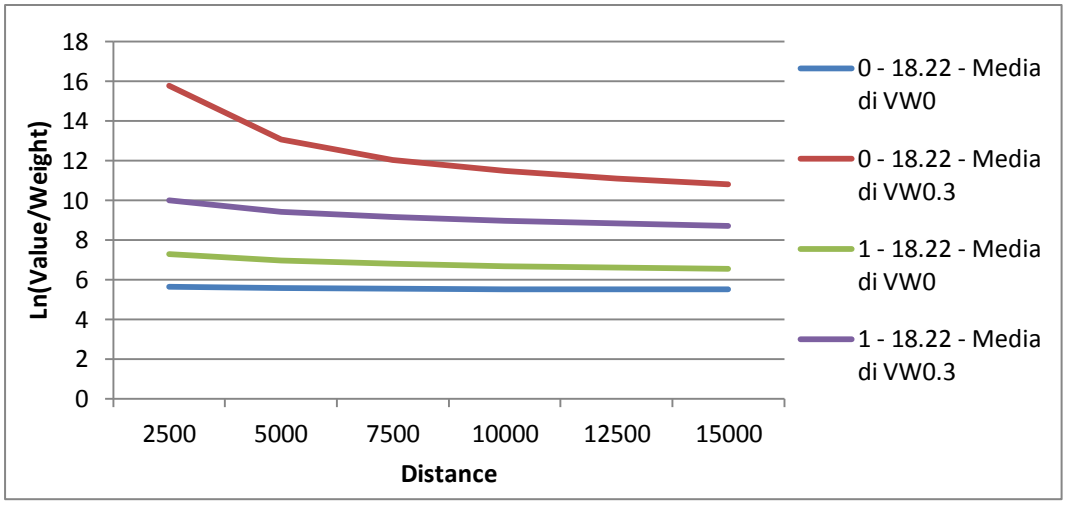


Figure 40 - ln(Thresholds V/W)=f(Distance) --> for Y=0 and Y=0.3 (Medium Fare3)

Third Fare = $e^{20.53}$ (Third Quartile)

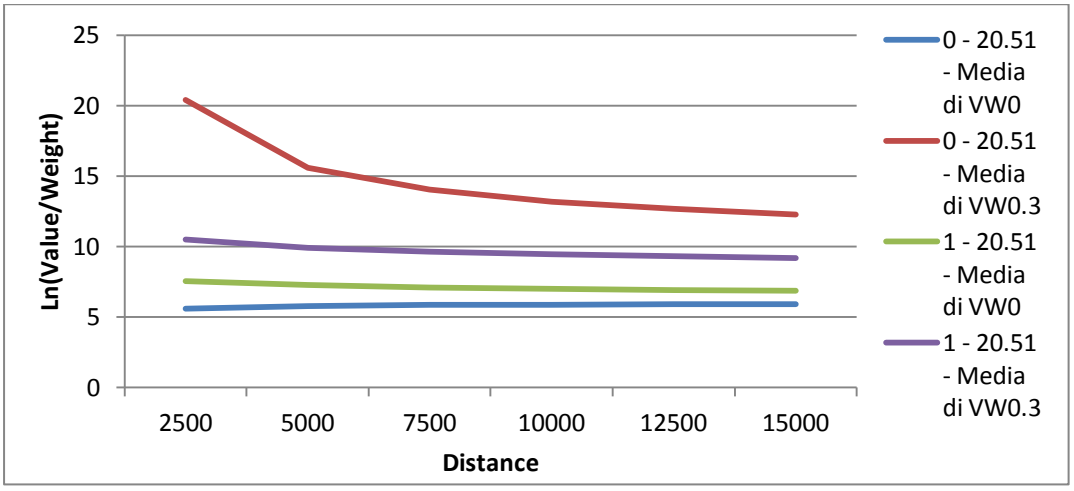


Figure 41 - ln(Thresholds V/W)=f(Distance) --> for Y=0 and Y=0.3 (High Fare3)

Setting Y=0.3, the minimum prices (value/weights) are much higher. We had to use logarithmic value/weights on the ordinate axis to be able to see the lines together since the slacks are huge. All of the 0.3 lines are decreasing monotonically. Even for the high fares lines (FareDifference3= $e^{20.53}$), minimum

value/weights are decreasing in distance. So, the anomalous trend we underlined before, related to the high fare level without time constraints is not consistent for high probabilities of using airplanes.

“Timely lines” result now under “non-timely lines”. Despite the much higher Value/Weight on average, timeliness constraint allows for a bigger price range to be traded “systematically” by air (y=0.3).

Range of goods commonly traded by planes (Y=0.3) is bigger if distance is bigger (for any volume, and density). It means that they start using systematically airplanes even for cheaper products if distance is bigger. The relative effect of cost that vary with distance (fuel, human resources) is even less important if we consider systematic usage of planes (Y=0.3), rather than the scenario in which “traders” start thinking about using airplanes in case of urgency (Y>0).

Obviously, the fare level strongly undermines the possibility to transport systematically by air, since it’s very different to sustain extra-costs once in a year (because of urgency) rather than to sustain these extra-costs any time company has to trade. Let’s have a look at the absolute value/weight values (instead of logarithms).

Low Fare scenario vs Median Fare scenario

T=0 (Non-Timely Trade)

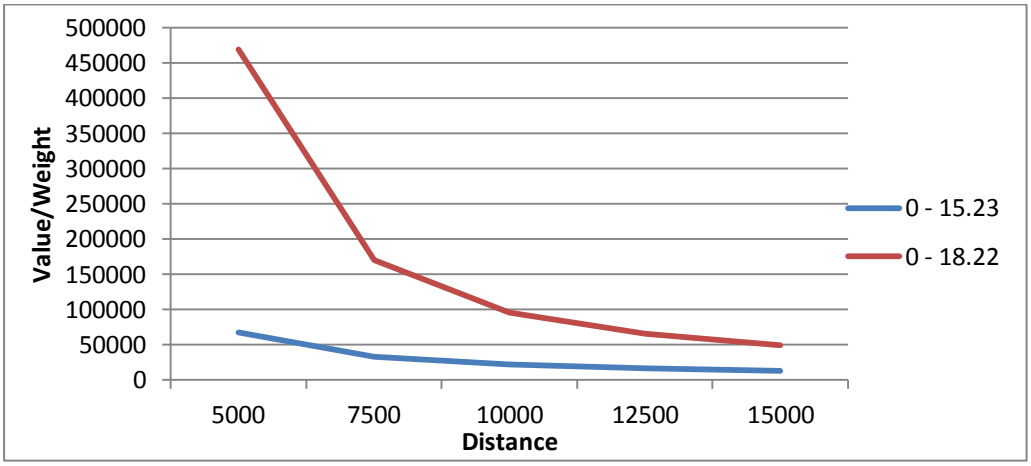


Figure 42 - V/W = f(Distance) --> "Non-Timely Scenario"

T=1 (Timely Trade)

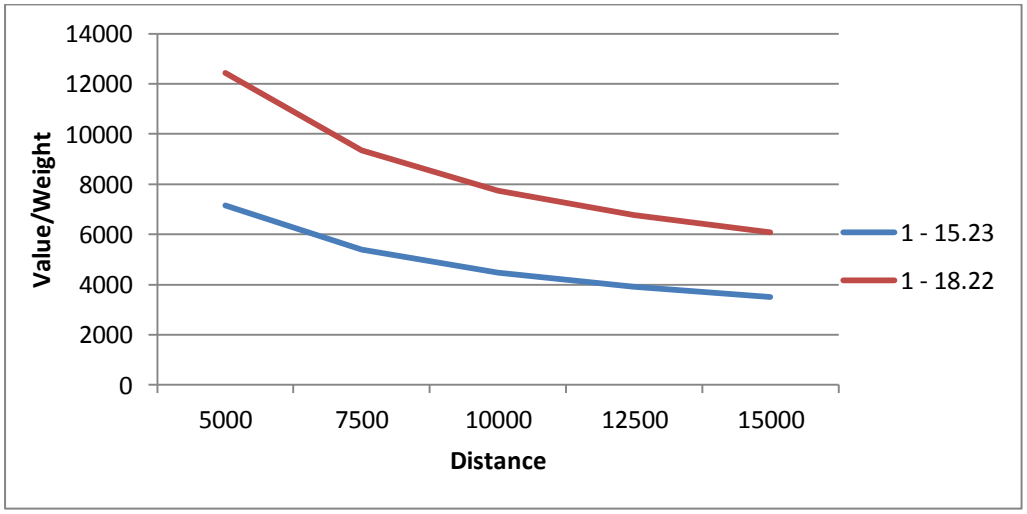


Figure 43 - V/W = f (Distance) --> "Timely Scenario"

If volume of trade and density are higher (Median Fare), extra-costs explode and huge future returns are needed to sustain costs.

We didn't consider the high fare in previous graph, because value/weights are too much bigger to be considered concurrently. We represent them by themselves here below.

High Fare Without Time Constraint

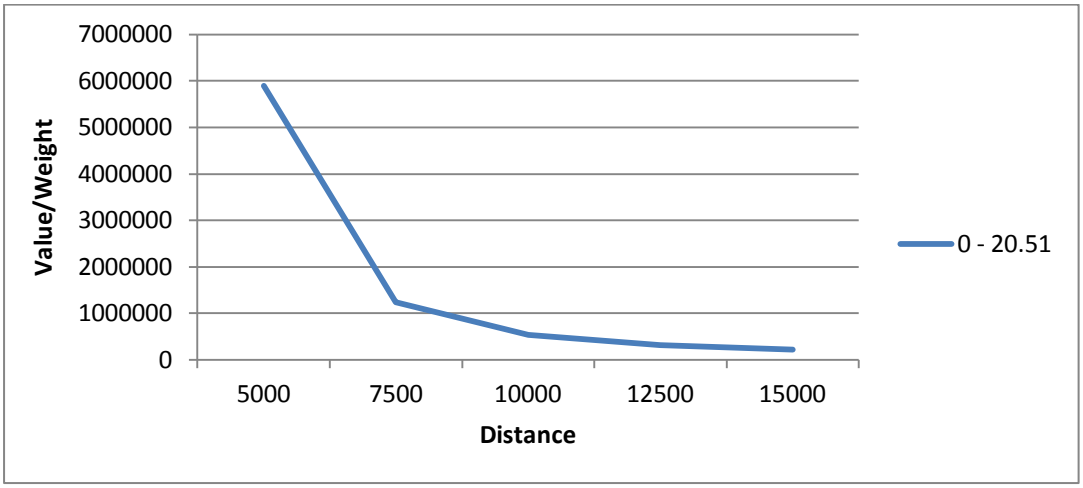


Figure 44 - V/W = f (Distance) --> "Non-Timely Scenario"

High Fare with Time Constraint

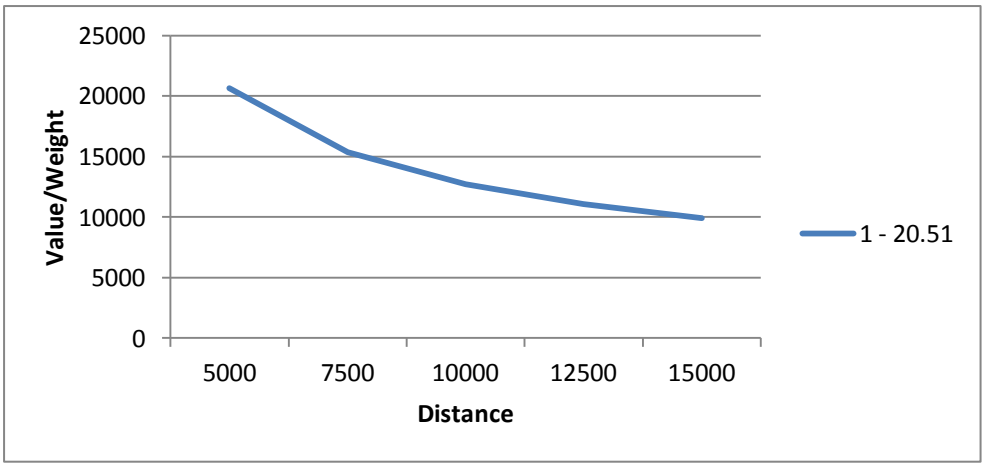


Figure 45 - V/W = f (Distance) --> "Timely Scenario"

Let's summarize the transportation patters identified for Y=30% in the table below:

	Low Fare		Median Fare		High Fare	
Distance	Timely 0	Timely 1	Timely 0	Timely 1	Timely 0	Timely 1
500	No Product	68799	No Product	134847	No Product	260625
1000	No Product	29958	No Product	55099	No Product	98538
2500	368025	12374	6996646	21818	No Product	36943
5000	67667	7151	469243	12441	5891362	20632
6000	48202	6277	286719	10901	2722330	18013
7500	33232	5389	169875	9344	1237399	15387
9000	25276	4782	116817	8288	717498	13616
10000	21815	4473	95844	7750	541272	12718
12000	17195	3997	69970	6926	348713	11349
12500	16346	3900	65497	6759	318374	11070
15000	13185	3503	49620	6073	218231	9937

Table 76 - V/W Thresholds Values (Y=0.3)

Where "No Product" is reported, it implies minimum value/weights higher than the highest value/weight in Eurostat Database. So, we assume there is no products traded systematically (Y=0.3) under these conditions.

The combined effect of "Time" in Framework

We justified results we got before by considering that higher distance implies on the one hand higher transportation costs (all of the applications based on iceberg models relies on this assumption), on the other hand strong time advantage. It results from analysis that a wider range of products is traded by air if distance is bigger and we conclude that the time advantage increases faster than extra-costs with distance in the nowadays scenario.

This assumption is intuitive and framework should stay substituting distance with time slacks related to distance in regression. We would expect similar results and critical points since we are now considering the variable for which there is direct causality relation with Y.

Here the equation identified by Stata:

$$Y = -57.5 + 9.89 * [\ln(V/W)] + 2.36 * [\ln(Fares\ Difference3)] - 53.55 * (Timeliness) + 0.19 * (Time\ Difference) - 0.39 * [\ln(V/W) * \ln(Fare\ Difference3)] - 0.037 * [\ln(Fare\ Difference3) * Time\ Difference] + 0.089 * [Time\ Difference * \ln(V/W)] + 0.047 * [(Time\ Difference) * Timeliness] + 8.63 * [\ln(V/W) * Timeliness] - 0.74 * [\ln(FareDifference3) * Timeliness]$$

Firstly, we represent functions on graph, to compare slopes.

While assessing effect of distance, we compared the 1000 km function with the 10000 Km one, distinguishing between the timely and non-timely trade. We will do the same now, considering the functions with time difference equal to -2 and 28 days. We choose these values since they could be adept at describing the 1000 and 10000 km distance.

"Non-Timely scenario" ("Normal products" non-time sensitive)

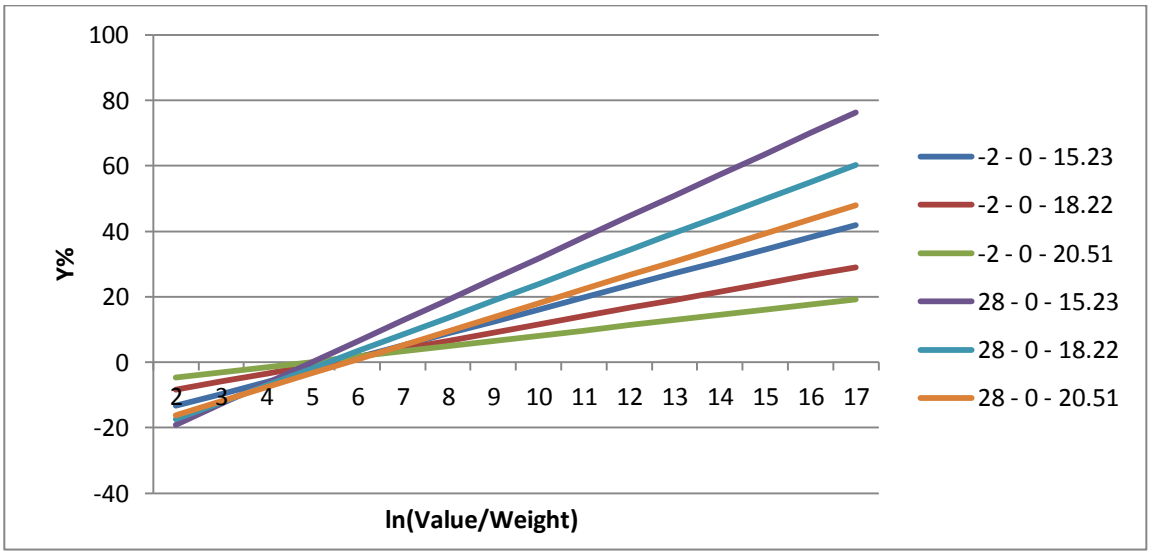


Figure 46 – Y%=f(ln V/W|Time Difference; Fare3) → "Non-Timely Scenario"

"Timely Scenario" (Products particularly time-sensitive)

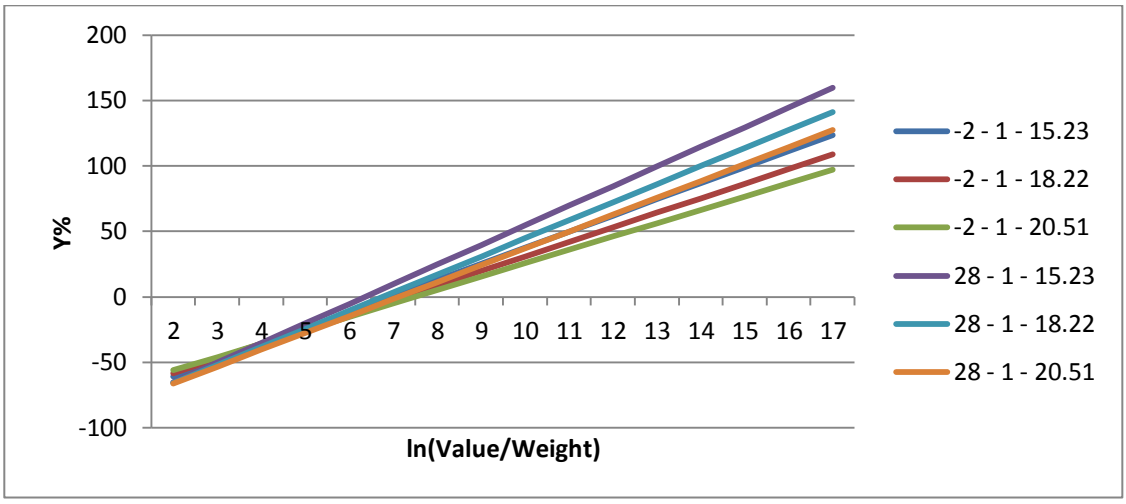


Figure 47 – Y%=f(ln V/W|Time Difference; Fare3) → "Timely Scenario"

We have same "lines sequence", and slopes of the lines are very similar to those we got using the distance variables. However, "slope slacks" look smaller here: there are lower ΔY here (passing from lines with the time difference variable equal to -2 days to lines with time difference 28) than before with distance (passing from 1000 Km to 10000 Km).

The next step is to compute the value/weight thresholds. Those values for which Y is equal to 0 (and later to 0.3) according to any possible "Time Difference" value.

We remind, setting Y=0 we find those value/weight that are "critical" since only one extra-euro implies positive Y (traders start thinking about planes in case of urgency); setting Y=0.3, we find those value/weight for which traders evaluate systematical usage of airplanes to deliver.

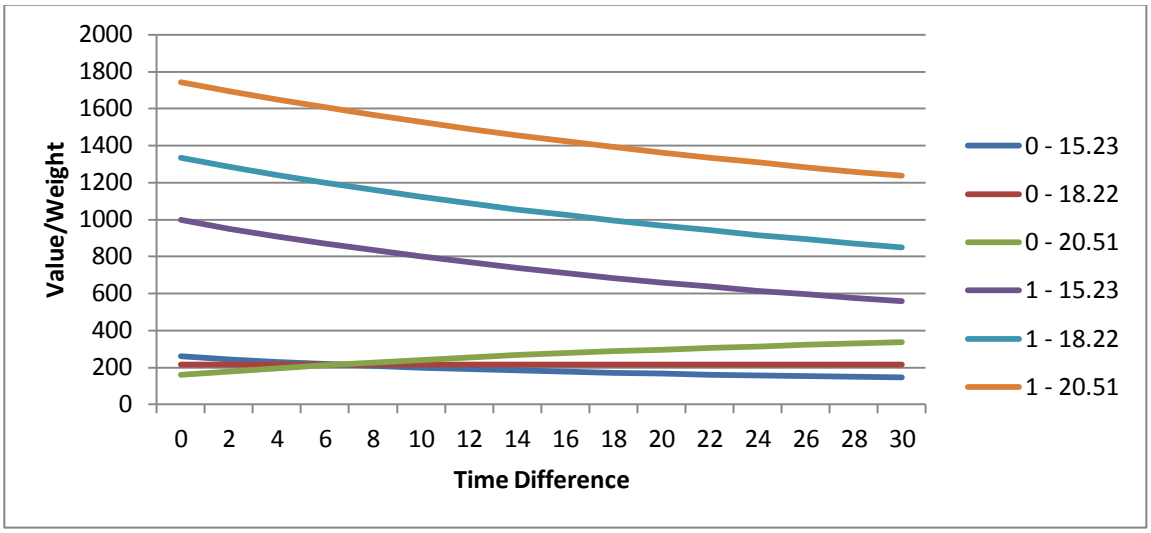


Figure 48 - Thresholds V/W =f (Time Difference) → Case:Y=0%

Curves are very similar to those we found with the "distance" variable in regression. However absolute values seem bigger.

In the Appendix G we report the graphs for the comparison of V/W in case of Y=0 and Y=0.3.

In the tables below there are the "values/weights thresholds" for both the cases Y=0 (table 77) and Y=30% (table78).

Y=0	Low Fares		Medium Fares		High Fares	
Days	**Timely 0**	**Timely 1**	**Timely 0**	**Timely 1**	**Timely 0**	**Timely 1**
-2	279	1046	217	1386	139	1796
0	261	997	217	1334	160	1744
2	245	952	217	1286	179	1696
4	231	910	217	1241	196	1650
6	219	871	217	1199	213	1606
8	209	834	217	1160	228	1566
10	200	800	216	1123	242	1527
12	191	769	216	1088	254	1491
14	184	739	216	1055	266	1456
16	178	711	216	1025	277	1424
18	172	685	216	995	288	1393
20	166	661	216	968	297	1363
22	161	637	216	942	306	1335
24	157	616	216	917	314	1309
26	153	595	216	893	322	1283
28	149	576	216	871	329	1259
30	146	558	216	850	336	1236

Table 77- Thresholds V/W --> Y=0

Y=0.3	Low Fares		Medium Fares		High Fares	
Days	**Timely 0**	**Timely 1**	**Timely 0**	**Timely 1**	**Timely 0**	**Timely 1**
-2	1301540	12376	No Product	21034	No Product	34248
0	795988	11359	No Product	19373	No Product	31601
2	509973	10452	10354229	17892	No Product	29241
4	340131	9641	5288918	16566	No Product	27129
6	234931	8913	2924589	15375	No Product	25235
8	167317	8259	1727630	14303	No Product	23530
10	122423	7669	1078897	13335	20750503	21990
12	91741	7135	706349	12459	10441390	20596
14	70225	6652	481539	11664	5719520	19331
16	54785	6213	339947	10941	3360419	18180
18	43475	5814	247387	10281	2093329	17130
20	35032	5449	184874	9678	1369990	16169
22	28623	5116	141422	9126	935074	15289
24	23683	4811	110442	8619	661665	14480
26	19820	4531	87846	8153	483032	13736
28	16761	4274	71027	7724	362331	13050
30	14311	4037	58279	7328	278332	12416

Table 78- Thresholds V/W --> Y=30%

With “Time Difference” we have same trends, but higher “Value/Weight” thresholds. The time difference indicator is less accurate since we took into consideration different transportation modes according to the different routes (airplanes vs trucks for Russia, airplanes vs ships for farer transoceanic routes). Indeed, the R^2 indicator passes from 0.39 with distance to 0.37 with time. The structure of the curve we drew proves that distance matters in our model especially from a time perspective. However, absolute values could be weird not only because of “low precision”: there can be some other causal effects implied by distance that we are not able to capture by substituting time.

Even for our model, the distance puzzle arises. It matters a lot, and influence transportation patterns but we have shown before, cost component influenced by distance is not the most relevant one (Fare Difference 1). Moreover, distance should be fully explained by time advantage caused by distance but results are not perfectly fitting. **According to literature, distance implies many other indirect consequences such as market uncertainties, informational problem, cultural barriers, infrastructure availability (for some zones or routes) and other countries’ features.**

5.3 Robustness Analysis

In order to verify the validity and the robustness of the previous regression models, we implemented another type of regression in which the dependent variable is binary.

Since the dependent variable is a probability, it's reasonable to adopt a model that constraints the predicted values to assume values "0" or "1": in this way we overcome the previous problem of having insignificant negative values!

We introduced the **Probit** model that is a particular regression whose coefficients are estimated according to the "*Maximum-Likelihood Estimation*" (MLE) method.

The Probit model is defined in the following way:

$$\mathbf{Pr}\,(\boldsymbol{Y} = \mathbf{1}|\boldsymbol{X}) = \boldsymbol{\Phi}(\boldsymbol{\beta_0} + \boldsymbol{\beta_1 X})$$

where $\boldsymbol{\Phi}$ is the cumulative distribution function of a Standard Normal.

The limit of this model is that coefficients are not directly interpretable. Positive coefficients indicates that if the variable goes up, there is higher probability of having Y=1. However, we cannot say anything about the coefficient value. The solution is to compute ΔY making the variable vary (ΔY=f(ΔX)).

In probit regression, R^2 is not a good indicator of model fitting. STATA gives in output the so called Pseudo-R^2; it's a corrected R^2 that measures the adaptation level of the model by using the maximum-likelihood function.

Going back to our case, we transform the independent variable in a binary one: we assign "1" when the probability of selecting "air shipping" is higher than 30%, "0" for all the other observations.

In table 79 we report the regression results. In regression (1) we introduce just the main variables without considering any interdependency. The signs of each variable are coherent (the sign is correct) and the most of the variables are significant.

Regressions (2) and (3) include all the interdependences among variables: in (2) we consider "Distance" while in regression (3) we substitute "Distance" with "Time Difference". In both of the cases we get quite good Pseudo-R^2. However, many interdependency coefficients are not very significant in regression (3).

REGRESSOR		**Probit** (1)	**Probit** (2)	**Probit** (3)
ln Value/Weight	Coef.	.4115996	-.7020527	-.6065093
	SE	.0355984	.3492529	.1632497
	z	11.56	-2.01	-3.72
Ln Distance	Coef.	.4624855	1.030819	
	SE	.0936227	.5301609	
	z	4.94	1.94	
Time Difference	Coef.			.0660963
	SE			.0307655
	z			2.15
Ln Fares3	Coef.	-.1149361	-.2101225	-.6468214
	SE	.0149487	.2313087	.1009747
	z	-7.69	-0.91	-6.41
Airport Density	Coef.	.0220811	.0220049	.0176415
	SE	.0147435	.0157304	.0157414
	z	1.50	1.40	1.12
Sq Airport density	Coef.	-.0005236	-.0006299	-.0008636
	SE	.0004041	.000448	.0004048
	z	-1.30	-1.41	-2.13
Time Sensitiveness	Coef.	.5751499	-6.706104	-1.092763
	SE	.1371959	2.873137	1.368413
	z	4.19	-2.33	-0.80
GDP	Coef.	1.46e-08	1.67e-08	2.18e-08
	SE	1.22e-08	1.29e-08	1.28e-08
	z	1.20	1.30	1.70
Fragility&Dangerous	Coef.	.7579386	.7833878	.7788915
	SE	.1375693	.1490942	.1476411
	z	5.51	5.25	5.28
Int Fare 3-Value/weight	Coef.		.0658561	.0694436
	SE		.0116453	.0114971
	z		5.66	6.04
Int Distance – Value Weight	Coef.		.0153615	
	SE		.0392639	
	z		0.39	
Int Time Difference – Value Weight	Coef.			-.0012202
	SE			.0028067
	z			-0.43
Int Fares3 – Timeliness	Coef.		-.02445	-.0446303
	SE		.0470913	.0464192
	z		-0.52	-0.96
Int Distance-Timeliness	Coef.		.5847192	
	SE		.2464246	
	z		2.37	
Int Time Difference - Value Weight	Coef.			.0166327
	SE			.013207
	z			1.26
Int Value/Weight – Timeliness	Coef.		.270662	.221882
	SE		.1376463	.1249308
	z		1.97	1.78
Int Fare3 - Distance	Coef.		-.0515661	
	SE		.0238377	
	z		-2.16	
Int Time Difference – Fare 3	Coef.			-.0025474
	SE			.0013076
	z			-1.95
Pseudo R^2		**0.3722**	**0.4159**	**0.4043**

Table 79 - Probit Regressions

6. Conclusions

We started our thesis aiming at understanding if the determinants identified by literature and interviews are actually able to describe the transportation choices in the nowadays context. Our goal was to give a contribution to the several questions that are largely debated in the most of the current economic literature about international trade patterns; in particular, we aimed at obtaining some empirical evidences about the current role of distance (*"Has distance died?"),* time (*"What is the role of time in the current globalized context?"),* costs (*"What is the effective weight of transportation costs?")* and infrastructures *("Are the infrastructures relevant for international trade?"*).

We identified *products determinants* such as "price", "density", "fragility" and "time constraints", and *country determinants* like "distance", "time to deliver", "transportation costs", "level of infrastructure" and "GDP". We checked for the relative importance of these variables and for the effectiveness of the different indicators built to measure them through regression analysis that we based on European trade data. We discovered that decisional choices about international trade and transportation means are affected by the concurrent presence of all these factors. Moreover, we looked for eventual synergic effects of variables by inserting all of the significant interdependencies in regression.

We discovered that product's price, need for timeliness, distance and costs are the most relevant determinants. All of the interdependencies among them must be taken into consideration and in particular the synergy between price and product time sensitiveness. Studying this interdependency, we also found that if product price is low, an eventual time constraint lowers strongly the airplanes

usage, while if price is high, it causes a strong increase of air shipping likelihood. We found also that distance doesn't represent a constraint anymore from a monetary costs perspective while it is increasingly important from a time to market point of view. We jumped to this conclusion because range of products that can be traded with planes is on average wider for longer distance (time advantage increases faster than extra-cost with distance). Finally, we discovered that the level of infrastructure is not a very prominent factor in transportation pattern for our sample; however, we also noted that air shipping could be "hard" for those developing (and underdeveloped) countries with poor infrastructures (e.g. Morocco). The lower transportation costs are pushing towards the market globalization for those countries that can sustain adequate infrastructures investments; so that, the poorest countries that can't sustain infrastructures expenditures are "cut off" from the global competition.

Let's now discuss briefly about further results' implications that can be evinced from our studies.

The first evidence that we got is the presence of **strong interdependencies among variables**. Variables by themselves cannot totally explain the "air shipping" probability: the final decision about transportation mode is the **closing output of a complex decisional process that takes into account many different determinants with all of their interdependencies** (Figures 49 and 50). In other words, we can never say that "*all the goods with a certain value/weight are traded by air*": in this preposition we are neglecting other relevant variables that strongly influence the final decision. So, we cannot define a rigid and general rule about transportation pattern, but we have to consider a more structured preposition such as "*airplane is selected for products with a certain value/weight, density and time sensitiveness, in case of a certain distance, fare and infrastructure level*".

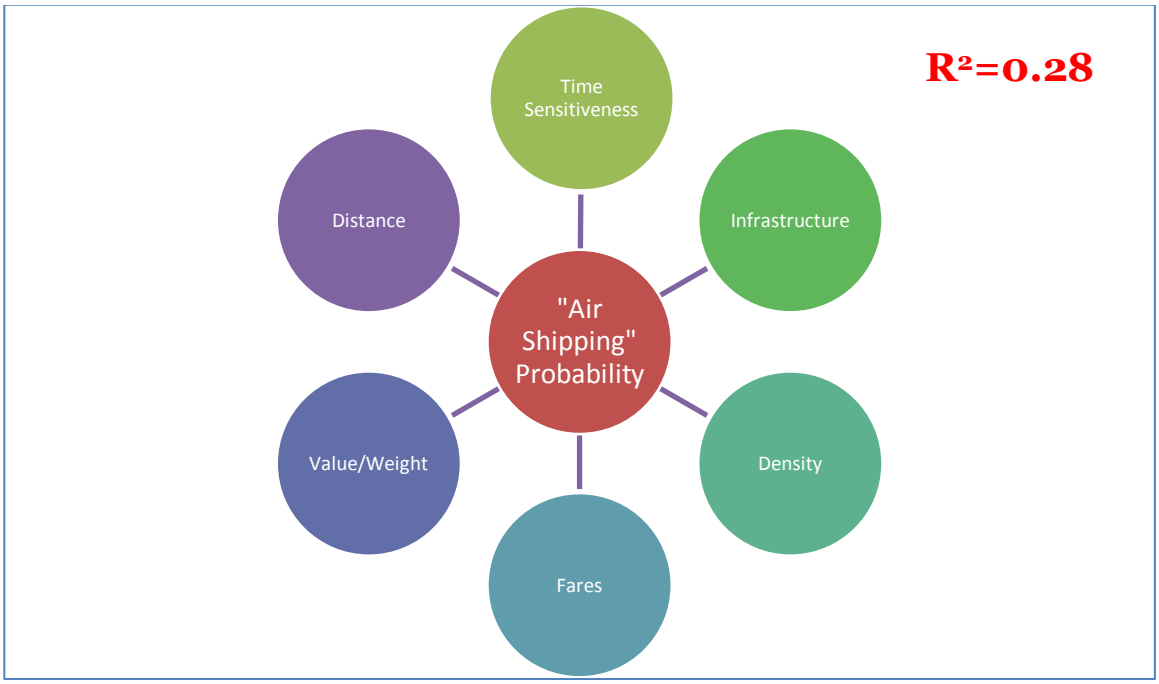


Figure 49 – Transportation Pattern Determinants

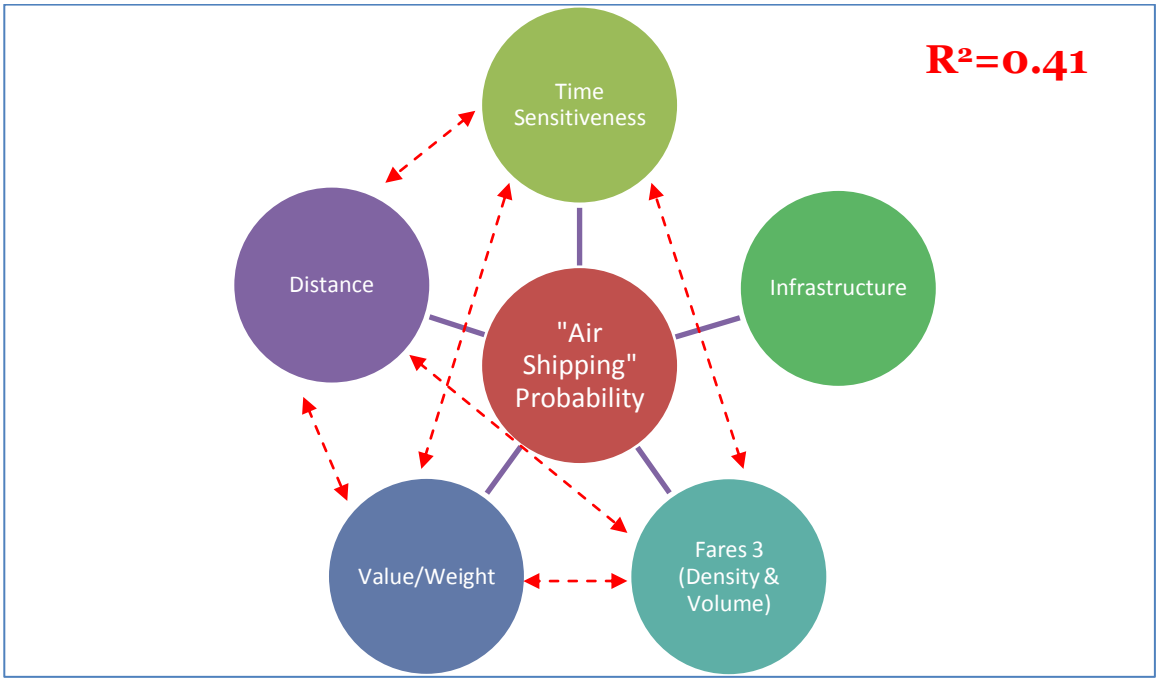


Figure 50 - Transportation Pattern Interdependent Determinants

Taking into account these important concepts, analyses that we carried out reveal many interesting results that match some of the most discussed economic topics of the current international context.

Several economists like Brun, Carrere, Guillaumont and Melo (2005) have discussed at length about the "puzzling" role of distance in international trade and many empirical evidences demonstrate that distance is still relevant in trade patterns despite the lower and lower transportation costs.

Our model confirms that distance has a strong importance in transportation patterns: the higher the distance the higher the probability of delivering by air. Through the analysis of the combined effect between distance and product value, we found that "air shipping" becomes theoretically unfeasible for any product if delivery distance is very low.

Even Hummels (2007) states that the strong reduction of transportation costs registered in the last decades should have caused an important decrease of trade sensitiveness to distance while in the reality, gravity models applications show that it keeps being an important constraint.

We investigated about real freight rates today, and we found that **fares not always reflect precise costs**. For instance, the interview with "Artsana S.p.A." revealed that fares per container to deliver in China are lower than fares to deliver in much closer countries in these months (despite transportation costs are obviously higher). This is because China needs empty containers to sustain its intensive export activities. So, demand and offer can change equilibrium prices.

Moreover, we discovered that those cost components depending only on distance are not influent in air shipping choice. On the contrary, all of the cost determinants related to cargo characteristics like density, unitary volume and trade volumes strongly affect the decisional process undermining the airplanes usage. So, even if transportation rates increase with distance, the biggest source of cost is related to the cargo nature. This could be evidence that reduction in

transportation costs doesn't imply any "Death of Distance": transportation costs decrease is pushing trade up but proportionally to distance (distance keeps being a constraint).

Moreover, distance is not "dying" because it can "absorb" the effect of many other related factors playing an "indirect role" in bilateral trade and transportation patterns. Economic literature is plenty of different discussions and theories about this. According to our studies, we can embrace the Harrigan(2005) theory about the **important role of time: nowadays, distance still matters because of "time implications"**.

Volatility and complexity of demand, shorter and shorter product life cycles and defragmentation of production processes around the globe reinforce the "need for speed and timeliness" making **the time dimension a critical success factor to exceed competitors**.

If a company is thinking about using airplanes, it's not a matter of distance itself but managers have to consider the **"premium price"** that firm can get through faster deliveries: **the higher the "time advantage", the higher the "air shipping" probability**. From our model, it's clear that "time advantage" is higher in case of "time sensitive" products: "air shipping" probability usually increases in case of "time constraints". This evidence matches completely the "*depreciation theory*" proposed by Hummels (2001) and Deardorff (2001) about the "role of time" in trade: products depreciate over time and "time advantage" derives from the fact that a "new" product is more valuable than an "older" one.

According to these results, we can uphold the repeated truism "**Time is Money**" that simply clarifies the higher and higher importance of time in the current international trade.

However, empirical outcomes evince that the traditional trade-off between "speed" and "trade cost" hasn't disappeared in trade patterns. "Timely deliveries" are always much more costly than slower ones and these **extra costs are sustainable only if "premium price" is high enough**. Decisional

processes about transportation modes must be always subject to the following constraint:

Premium Price > Extra Costs; With:

Premium Price = f (Value/Weight; Distance; Timeliness)

Extra Costs = f (Volume of Trade; Distance)

To better understand this trade-off, we computed all of the critical product values for which benefits exceed extra-costs, at the different distances, with or without time constraint. “Traders” can adopt air shipping systematically only for products in a very high price range; while products for which airplanes are supposed to be used only “sporadically” (in case or urgency) allow even lower prices.

It’s interesting to notice that “minimum prices” change with distance to be covered. Since minimum prices go down, the range of products for which airplane is affordable is wider for longer distances. It means that “time advantage” grows up faster than “extra-costs” on average.

An exception is constituted by those products that imply very high delivery frequency (e.g. wheat) and/or huge volumes of trade (e.g. coal). Resulting in very high transportation costs, these products show “minimum prices” that increase with distance. In this case, extra-costs grow faster than time advantage. However, this is true only for scenarios without time constraints and for commodities that are obviously shipped by air just in case of urgency. It’s enough to relax the first hypothesis considering time sensitive markets, and this result is not consistent anymore.

Once again, benefits of time depend on the importance that markets give to them.

So, our study reveals that **price of goods** plays a dominant role. In general, the higher the product value the greater the “air shipping” probability because “extra

costs" are more affordable and the "opportunity cost" is relevant for expensive products.

Particularly, airplanes' usage seems to be much more sensitive to price than to transportation costs and this trend is more marked for those products with time constraints.

The basic concept is that airplanes will never be selected if the "premium price" is not high enough to "sink" costs (even in case of time constraint) and customers have to buy from neighboring countries.

This point is directly connected to the well known concepts of "product substitutability" and "love for varieties". In the current context, customers want more and more personalized offers and they are willing to pay higher prices to get more varieties. As a consequence, the creation of highly differentiated product range makes the output more valuable but less substitutable. So, under this assumption "product substitutability" is inversely proportional to its price: the lower the price the higher the substitutability. We can coherently conclude that products for which "air shipping" is not affordable (for their low prices) are substitutable, so that customers can easily obtain them from neighboring countries.

Although distance and transportation costs have a significant role in our model, we have to stress that product value, time sensitiveness and especially their interaction are the strongest influencing factors.

We have discussed above about the role of distance and we said that time implications seem to be the most important reason why "distance isn't dying". Finally, we tried to verify this assumption by substituting distance with the correspondent "Time Difference" in regression.

We would have expected to replicate precisely the results we got before. Actually, we found very similar results, but the consistency of the statistical analysis looks lower. The distance puzzle, so well discussed in literature, arises again since we

are not able to fully understand the effect of distance in our model (even if time explains the largest part of it). However, we can reasonably list all of the factors discussed by authors that can be "hidden" behind the distance coefficient. Distance of trade can imply lack of information and asymmetries, different infrastructural levels, zones or countries effects, regionalism.

In other words, we are not able to capture some country features that can influence the role that distance is playing in regression (it's not a coincidence that results change by inserting some of the "country dummy variables"). Hence, the model is more adept at describing the European context, for which we collected data; further investigations referring to other zones in the world should be done for a more precise generalization.

Moreover, the indicators "Fare3" (the most used measure for transportation costs) and "Time Difference" can be somehow distort since we evaluated them comparing "air" and "road" features for close countries, while comparison regarded air shipping and sea shipping for trans-oceanic routes. The particular position of Switzerland, a country set within the boundaries of the EU27 (our point of reference), is difficult to be interpreted in analysis and it could have caused non-linear relationship for some determinants. Finally, our model is "static", since we are considering data referring to 2010 only. A comparison with past data could disclose some knowledge about future role of distance, transportation costs and potential for air shipping in trade.

In conclusion we can state that model results described above give a contribution for explaining some trends that characterize the current international trade. The increasing potential of the air shipping solution (higher need for time, lower air shipping costs) is actually changing the current international trade geography and composition.

Since the current context is characterized by a strong "love for variety" (customers are willing to pay higher prices to get more and more customized product varieties), companies should be able to produce a wider range of

products through the implementation of an ideal "*mass customized"* production. However, due to the concurrent presence of strong economies of scale and limited resources, companies have to specialize the production just in few sub-categories. Companies can sustain this kind of policy, since customers can get any product variety required from far countries by using the fast mean of transportation that is increasingly less expensive. **Under this perspective, lower transportation costs can be seen one of the enabling factors for the globalization (indeed, we discovered that time advantage grows faster than extra costs for longer distances)**. Economists would expect then a reduction of the trade sensitiveness to distance but as already said, this doesn't seem to be verified empirically. The point is that market differentiation brought some other concurrent trends: composition of trade can't be compared with the past anymore and commerce geography is different because even far countries can start being competitive in some given product sub-categories rather than others.

Moreover, the extreme differentiation brought to a concurrent counter-trend. If a given finished product has one month life-cycle, it means that time is more and more important in any phase of the supply chain since companies need raw materials to produce and market finished product quickly.

From our empirical analysis, we evince that finished manufacturing products can sustain air shipping for far distances, while raw materials and/or commodities are usually transported less by air in case of time constraint. We can deduce that more affordable transportation costs are enabling factors for market differentiation but an indirect consequence is the increasing need for speed at any level of the supply chain; this results in expensive products (air costs are "sinkable") to become absolutely **"not-substitutable"**: customers are willing to pay premium price to get the exact product variety they want from anywhere in the world; while cheap products (air costs are not "sinkable") such as raw materials and commodities become **"fully substitutable"** so that

companies will be content with the neighboring countries' offer. **The higher "close bilateral trade" for commodities and the intra-industry trade** (if there are many different product varieties, there can be trade between close countries specializing in the same given industry, but different "product types") **can be a reason why trade sensitiveness to distance does not decrease over the years in gravity model applications.**

This is evidence that transportation costs and technological breakthrough are actually radically changing international commerce.

Bibliography

REFERENCES

- **Anderson James E., van Wincoop Eric**. (September 2004) "Trade Costs". Journal of Economic Literature Vol. XLII, pp. 691-751

- **Baier Scott L., Bergstrand Jeffrey H.** (2001) "The growth of world trade: tariffs, transport costs, and income similarity". Journal of International Economics, No.53, pp.1 -27

- **Ballou Ronald H.**, "Logistics, Supply Chain and Transport Management", Cambridge International College.

- **Bayus Barry L.**, (1998)"An Analysis of Product Lifetimes in a Technologically Dynamic Industry", Management Science/Vol. 44, No. 6

- **Bernhofen Daniel, El-Sahli Zouheir, Kneller Richard.** (2011) "Containerization and International Trade. Initial Explorations." CESifo

- **Berthelon Matias, Freund Caroline**. (2008) "On the Conservation of Distance in International Trade". Journal of International Economics, No.75, pp.310-320

- **Bollen F., Kissling C., Emond J., Brecht J., McAneney J., Leake J., Crompton R., Nunes C., Metz A., Duval K., Laniel M., Ye J.** (2004) "Sea and Air Container Track and Trace Technologies: Analysis and Case Studies", Project No. TPT01/2002T.

- **Brun Jean-Francois, Carrère Céline, Guillaumont Patrick, de Melo Jaime.** (2005) **"**Has Distance Died? Evidence from a Panel Gravity Model**".** The World Bank Economical Review.

- **Cachon Gérard P., Swinney Robert**, (2011) "The Value of Fast Fashion: Quick Response, Enhanced Design, and Strategic Consumer Behavior", Management Science, Vol. 57, No. 4, pp. 778–795

- **Coe D. T., Tamirisa N. T., Bhavnani R.** (2002) "The Missing Globalization Puzzle" IMF Working Paper 02/171, International Monetary Founds

- **Commission of the European Communities**, (2006) "Freight Transport Logistics in Europe – the key to sustainable mobility", COM 336

- **Confederazione generale italiana dei trasporti e della logistica**, (1997) "Il Trasporto Marittimo"

- **Coughlin C. C.** (2004)"The Increasing Importance of Proximity for Exports from U.S. States. Federal Reserve Bank of St. Louis Review 86, 1-18.

- **De Gucht K.** (2010)"A renewed trade strategy for Europe: Speaking points", Presentation of the "Trade, Growth and World Affairs" Communication, European Commission, Brussels

- **Deardorff Alan V**. (2001) “Time and Trade: The Role of Time in Determining the Structure and Effects of International Trade, with an Application to Japan”. Paper prepared for the conference, “Analytical Issues in the Trade, Foreign Direct Investment, and Macro/Financial Relations of the United States and Japan”.

- **Dias D.A.**. (2010) “Gravity and Globalization”, mimeo UCLA

- **Disdier Anne-Célia, Head Keith**. (2008) “The Puzzling Persistence of the Distance Effect on Bilateral Trade”. The Review of Economics and Statistics, Vol. 90, No. 1, pp.37-41.

- **Djankow, Freund and Pham.** (2006) “Trading on time”. World Bank Working Paper, No. 3909

- **Duranton G., Storper M.** (2008)“Rising Trade Costs? Agglomeration and Trade with Endogenous Transaction Costs”. Canadian Journal of Economics, 41(1), 291-319

- **Estevadeordal Antoni, Frantz Brian, Taylor Alan M**. (2003) “The Rise and Fall of World Trade, 1870–1939”. Quarterly Journal Of Economics, Vol.118, No.2

- **European Commission for Trade.** (2010) "Trade, Growth and World Affairs"

- **Evans Carolyn L., Harrigan James.** (2003)“Distance, Time, and Specialization”. National Bureau of Economic Research. (NBER Working Paper 9729)

- **Feenstra Robert C**., (1998) “Integration of Trade and Disintegration of Production in the Global Economy”. Journal of Economic Perspectives. Vol.12, No. 4, pp.31-50

- **Ford Henry.** (1922) “My Life and Work”, chapter IV, pages 72-73

- **Grossman G.M.** (1998) “The Regionalization of the World Economy. University of Chicago for the NBER.

- **Harrigan James, Venables Anthony J.** (2004) “Timeliness, Trade and Agglomeration”. National Bureau of Economic Research. (NBER Working Paper 10404)

- **Harrigan James**. (2005) “Airplanes and Comparative Advantage”. National Bureau of Economic Research. (NBER Working Paper 11688)

- **Hayakawa, Kazunobu,** (2008) “The Choice of Transport Mode: Evidence from Japanese Exports to East Asia”, IDE Discussion Paper. No. 155

- **Hornok Cecília**. (2009) “Need for Speed: The Role of Timeliness in the Trade Effect of EU Accession”.

- **Hummels David.** (2001) **“**Time as a Trade Barrier”. GTAP Working Paper No. 18

- **Hummels David**. (2007) “Transportation Costs and International Trade in the Second Era of Globalization”. Journal of Economics Perspectives ,Volume 21, Number 3, pp. 131-154

- **International Chamber of Shipping.** (2007), Sofia Ministerial Conference on Transport Congestion

- **International Transport Forum**. (2010) OECD department, "Moving Freight with Better Trucks"

- **Kapoor, Kansal**. (2005) "Basics of Distribution Management – A Logistical Approach", p.89

- **Kotler Philip.** (1989) **"**From Mass Marketing to Mass Customization", Planning Review, pp. 47

- **Krugman Paul**. (1995) "Growing World Trade: Causes and Consequences", Stanford University, Brookings Papers on Economic Activity, Vol. 26, No.1, pp. 327-377.

- **Krugman Paul**, (2009) "Citigroup Foundation Special Lecture, Festschrift paper in Honor of Alan V.Deardorff", University of Michigan IPC working paper 91.

- **Li Yue, Wilson John S.** (2009) "Time as a Determinant of Comparative Advantage". Policy Research Working Paper 5128, Development Research Group (World Bank)

- **Leamer E.E., Levinsohn J.** (1995) "International Trade Theory: The evidence". In: Grossman G., Rogoff K., Handbook of International Economics, vol. 3, pp 1339-1394.

- **Leamer E.E., Medberry C.J.,** (1993) "U.S. Manufacturing and an Emerging Mexico", NBER working paper 4331, National Bureau of Economic Research, Inc.

- **McCallum J.** (1995) “National Borders Matter: Canada-U.S. Regional Trade Patterns”, American Economic Review, 85(3), pp.615-623

- **Micco A., Serebrisky T.** (2006) “Competiona regimes and air transport costs: The effects of Open Skies Agreements”, Journal of International Economics, Issue 70, pp.25-51

- **Noblet, Sandrine and Belgodere, Antoine.** (2010) “Coordination cost and the distancepuzzle”. MPRA Paper, No. 27502

- **Nordås Hildegunn Kyvik, Pinali Enrico and Geloso Grosso Massimo,** (2006)“Logistics and Time as a Trade Barrier”, OECD trade policy working paper no. 35

- **Regolamento (CE) N. 451/2008 DEL PARLAMENTO EUROPEO E DEL CONSIGLIO** del 23 aprile 2008 che definisce una nuova classificazione statistica dei prodotti associata alle attività (CPA) e abroga il regolamento (CEE) n. 3696/93 del Consiglio

- **Tanaka Kiyoyasu**, “Transportation Cost and Distance: Evidence from the Census of Logistics”, 2009

- **The World Bank Group**, (2009) “Air Freight: A Market Study with Implications for Landlocked Countries”, Transport Papers 26

- **Tsai Ying Yi**, (2009) “Strategic Choice of Freight Mode and Investments in Transport Infrastructure within Production Networks”, ADBI Working Paper Series, No. 132

- **UNDP (United Nations Development Programme)**, (2008) "Shipping and Incoterms – Practice Guide"

- **Van de Reyd J.K.R., Wouters M.V.L.J.,** (2005) "Air Cargo Density Research", Delft University of Technology

BOOKS, MANUALS and CLASSES SLIDES

- "In Defence of Globalization", Bhagwati Jagdish**,** 2004, Oxford University Press.
- "International Business. Economia e strategia internazionale: l'impresa nei mercati globali", Charles W.L.Hill, 2007 (Hoepli)
- "Introduzione all'econometria", James H. Stock, Mark W. Watson, 2009 (Pearson)
- "Just in time", D. Hutchins, Gower technical press, 1988. Number of pages: 215. O'connor, P. D. T. (1989). Quality and Reliability Engineering International, 5: 183.
- Professor L.Tajoli, "International Economics" lectures (a.y.2010)
- Professor M.Taish, "Industrial Technologies" lectures slides (a.y. 2010)
- Professor V.Gekara, "International Trade Operations" lectures slides (a.y. 2009) – School of Management – RMIT (Royal Melbourne Institute of Technologies)
- Professors A.Perego, R.Mangiaracina "Logistics Management" lectures slides (a.y. 2010)
- Professors G.Noci, L.Lamberti "Marketing" lectures slides (a.y. 2010)

SOURCES OF DATA

- "U.S. Department of State": http://www.state.gov/e/eeb/rls/fs/2009/119760.htm
- Airbus: http://www.airbus.com/
- Artsana S.p.A.: http://www.artsana.com/
- Airports Council International http://www.aci.aero/cda/aci_common/display/main/aci_content07_banners.jsp?zn=aci&cp=1_725_2__
- Boeing: http://www.boeing.com/commercial/cargo/01_05.html
- CEPII: http://www.cepii.fr/
- Delivery Time simulation: http://www.tnt.it/it/itsolutions/tempi_consegna.shtml
- DHL: http://wap.dhl.com/serv/volweight.html
- EUROSTAT: http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database
- FedEX: https://www.fedex.com/ratefinder/home?cc=US&language=en&locId=express
- Flightstat: http://www.flightstats.com/go/Airport/airportsOfTheWorld.do
- GoogleMaps: http://maps.google.it/
- International transport fourum: http://www.internationaltransportforum.org/
- IQ Global: http://www.iqglobal.com/
- ITALMONDO : http://www.italmondo.com/it/index.html
- Maritime routes calculation: http://www.searates.com/
- NATIONMASTER: http://www.nationmaster.com/graph/tra_roa_pav_of_tot_roa-transportation-roads-paved-of-total

- OECD: http://www.oecd.org/home/0,2987,en_2649_201185_1_1_1_1_1,00.html
- Oxford Dictionary: http://oxforddictionaries.com/
- Data about Specific Weight: http://www.asiinstr.com/technical/Material_Bulk_Density_Chart_T.htm
- The Geography of Transport Systems: http://people.hofstra.edu/geotrans/index.html
- UN comtrade: http://comtrade.un.org/
- UNCTAstat:http://unctadstat.unctad.org/ReportFolders/reportFolders.aspx?sCS_referer=&sCS_ChosenLang=en
- UPS:http://www.ups.com/content/it/it/contact/index.html?WT.mc_id=TMP_Italy_37917&WT.srch=1
- VALOCCHI: http://www.valocchi.eu/logistica/trasporti.htm
- ViaMichelin: http://www.viamichelin.it/
- WORLD BANK: http://www.worldbank.org/
- World Economic forum: www.**weforum**.org
- WTO: http://www.wto.org/index.htm

Appendixes

APPENDIX A

Economic Models

Some economists built econometric models to understand how globalization and distance changed the international trade. In the Eaton and Kortum model (2002), trade costs are still considered in the "iceberg form": to deliver one unit from country i to country n, it requires shipping $D_{ni} > 1$ units.

The price of a good j produced in country i and bought in country n is:

$$P_{ni}(j) = [C_i/Z_i(j)]^* D_{ni}$$

Where C_i is the cost of inputs to produce j, Z_i discounts the cost while considering the technological efficiency of the producer country. D_{ni} (distance between country n and country i) adjusts prices considering "melting of value during shipment". Eaton and Kortum go through a complex probability analysis to demonstrate numerically that trading with more countries and technological improvements lower prices, while greater distance increases prices.

Harrigan (2005) "adjusts the EK model" considering the need for "timeliness".

1- He adds a "timeliness factor" in a Cobb-Douglas preferences function: "*consumers prefers fast delivery, if marginal contribute of time exceeds premium price for fast delivery*".

f(z)>P(z)/p(z) where P(z) is the price for fast delivery and p(z) the price for common delivery.

2- He defines the "supply side": "in a perfect competitive system, FOB price is the exact unit cost, and CIF price comes adding only transportation costs". Shipping costs are in the iceberg form (it's exactly as you have to deliver something more than one unit to make one unit arrive).
T(z) are shipping costs, and it can be equal to s(z) or a(z) according to the timeliness preference. A(z) is the cost for slow delivery, while S(z) the cost for fast delivery.

*a(z)/[s(z)*f(z)]>0*

Note that the f(z) has been already used to understand the utility function at point 1.

3- **The shipping cost is adjusted by considering "timeliness discount"**. If the delivery is timely, you will have a lower T(z).

T(z)= min [s(z);a(z)/f(z)] → *for far country*
T(z)= s(z)/f(z) → *for close country (because you can only be fast)*

And, as in EK model:

FOB price= $W_o/B_o(z)$
CIF price= $P_{od}=[W_o*T_{od}(z)]/b_o(z)$

Where W_o is the cost labor in origin country, and the productivity variable of the origin country B_o is seen as a "Frechet distribution" (random casual variable). You are finally able to determine the probability for a given

producer country to get a reference destination country according to the relative price parameter (you get it if your price is the lowest one).

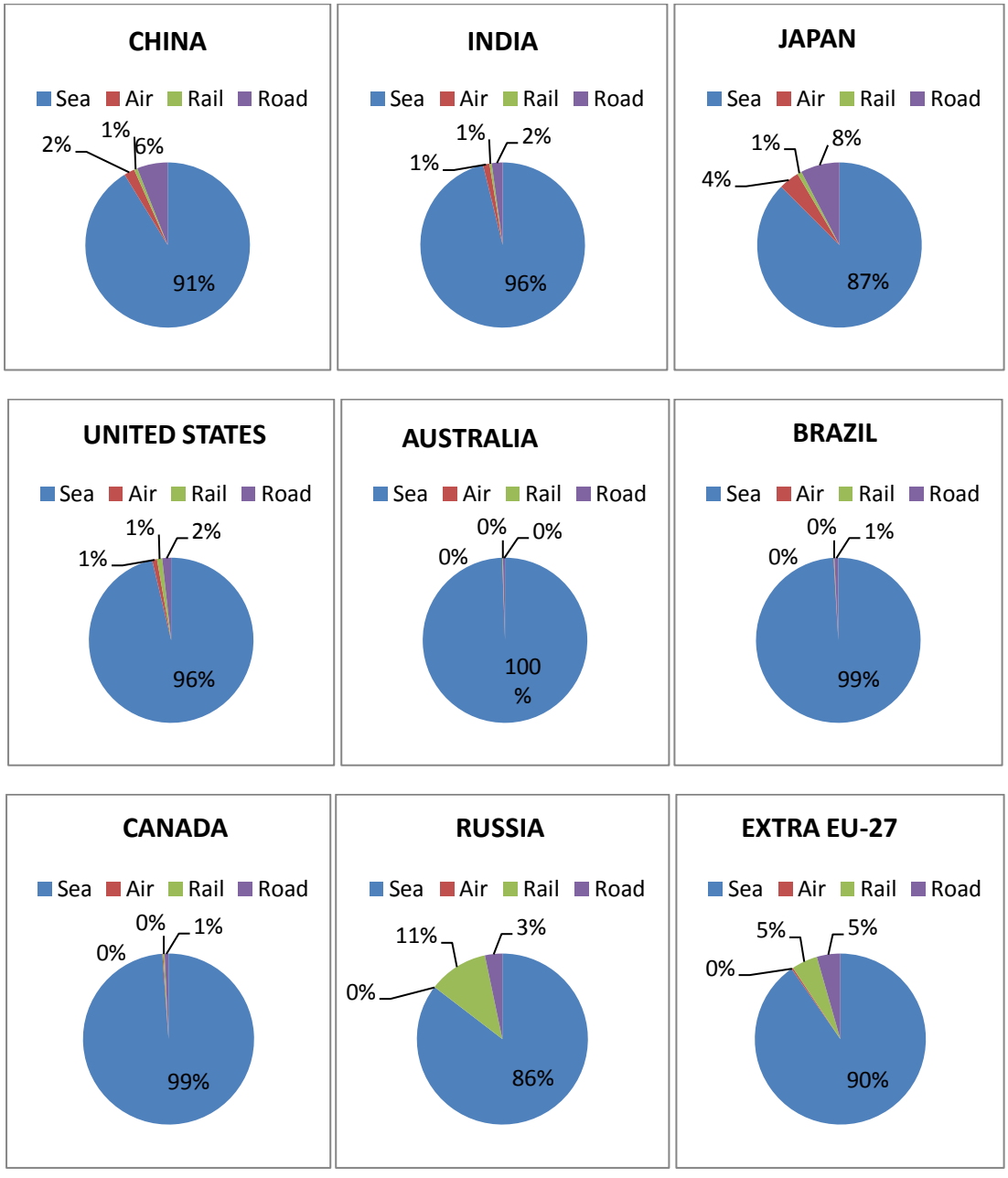
4- Like in the "EK model", Harrigan defines expected prices, national incomes and welfare: if the need for timeliness represented by f(z) increase, average prices increase.

Then he defines a new "comparative advantage model" considering distance, timeliness and cost for different transportation means. The conclusion of the theoretical model is that: "*faraway countries have comparative advantage in light products (referring to value weight ratio), while close countries in heavy products, that is, distant countries will have relatively higher market share in goods shipped by air*".

APPENDIX B

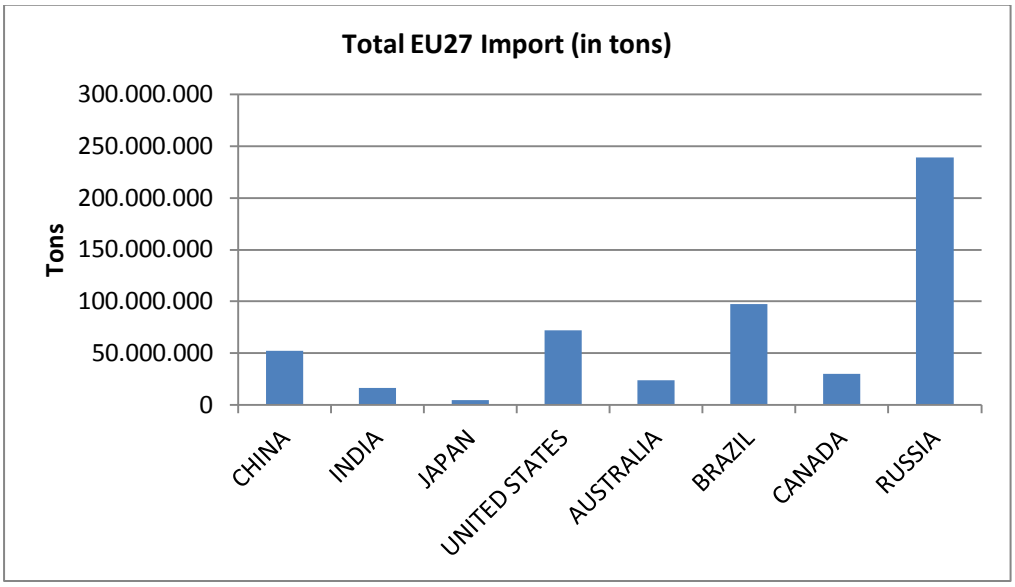
Focus on European Trade

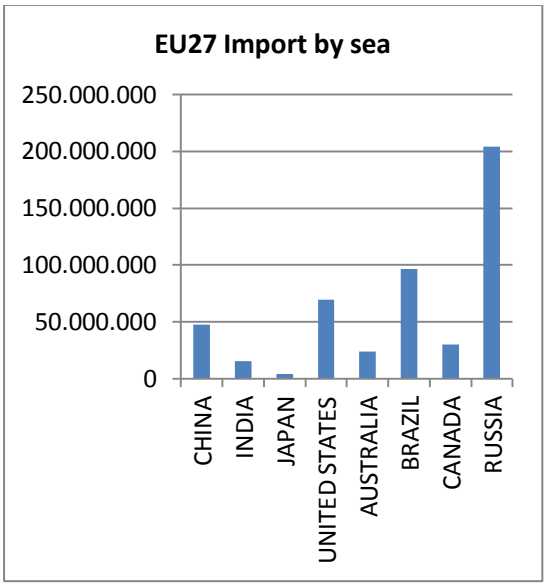
Weight Perspective→ % of trade in weight (tons) by mode of transport.

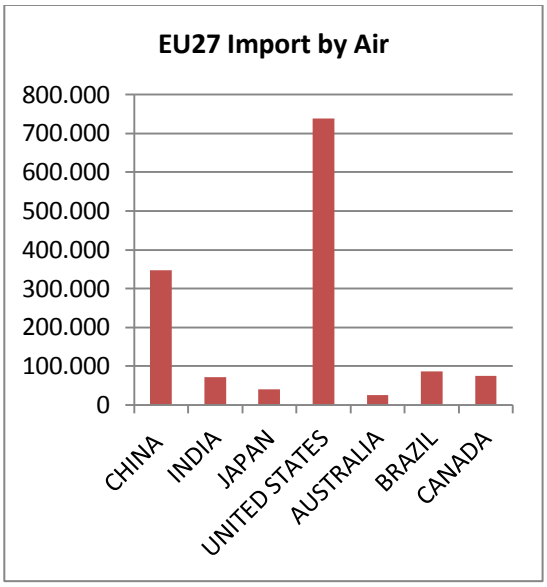


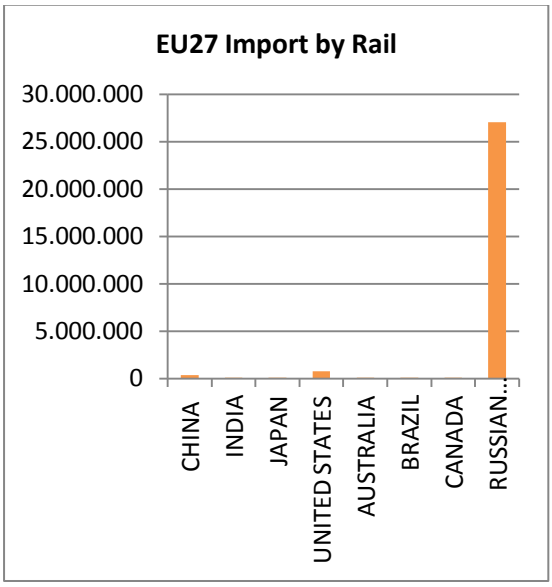
European Trade Partner in tons:

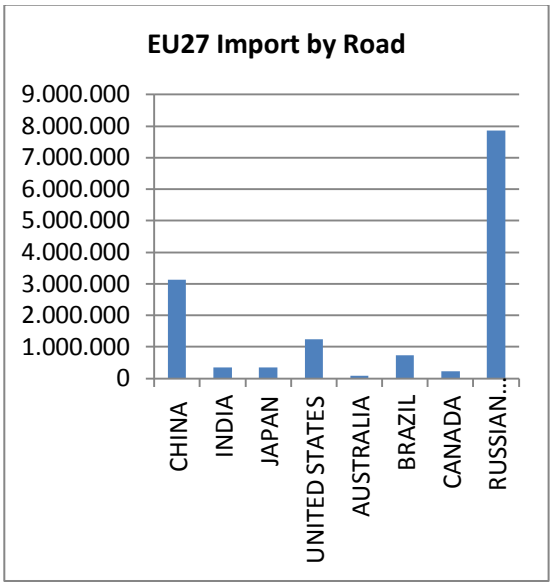
Total EU27 import (sea + air + rail + road) in tons.





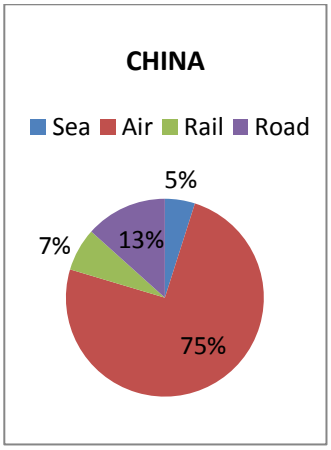


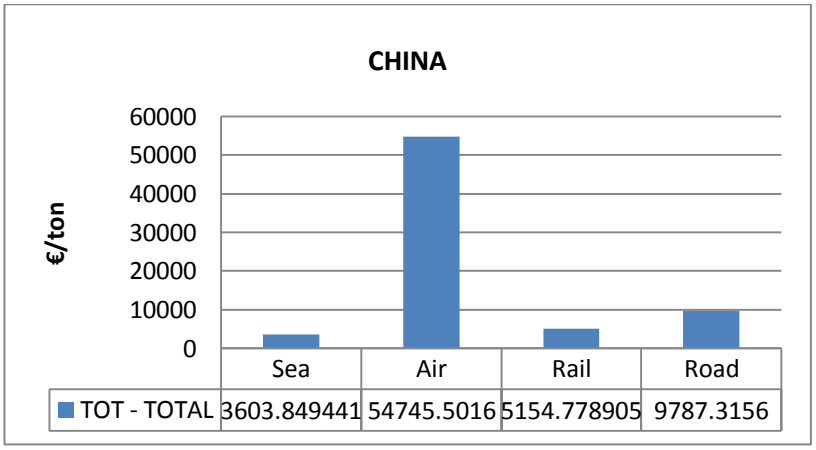


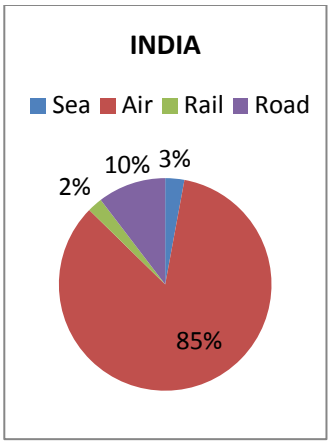


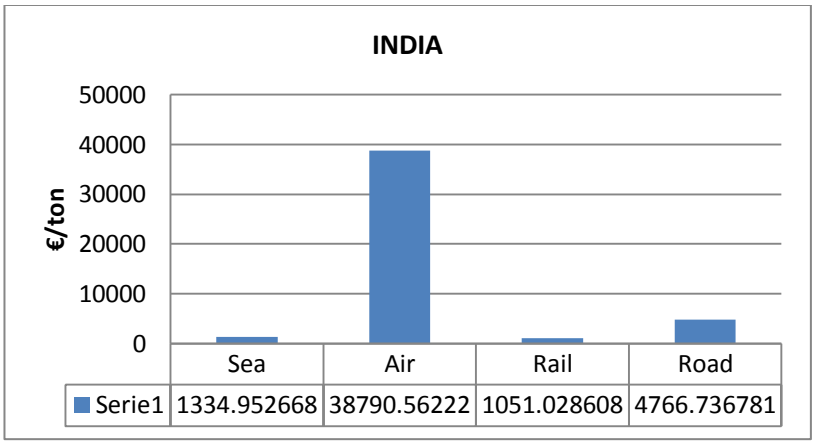
Value/Weight Analysis:

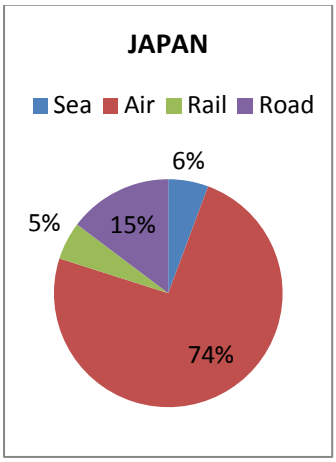
We report the results of the value/weight analysis for each country-partner.

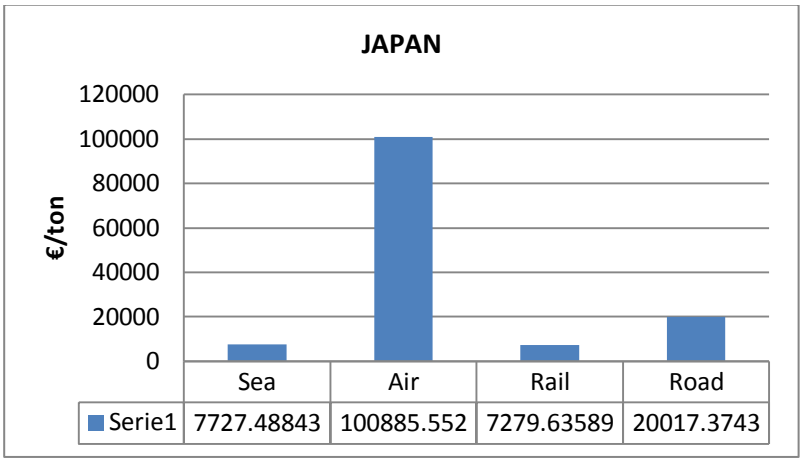


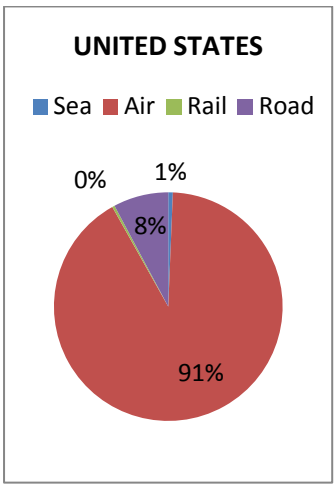


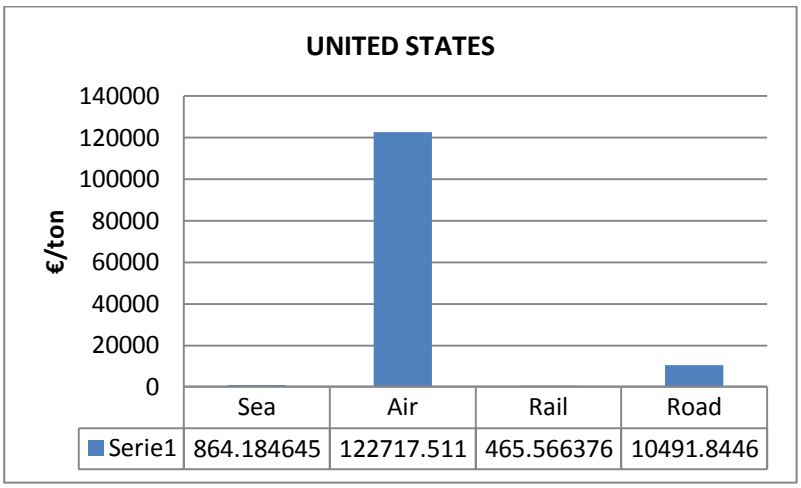


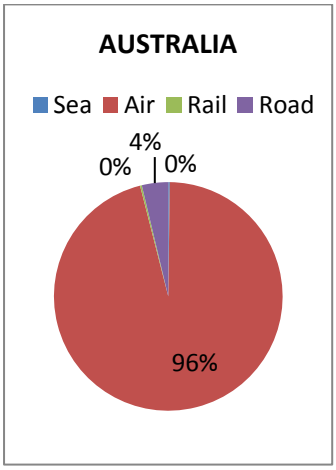


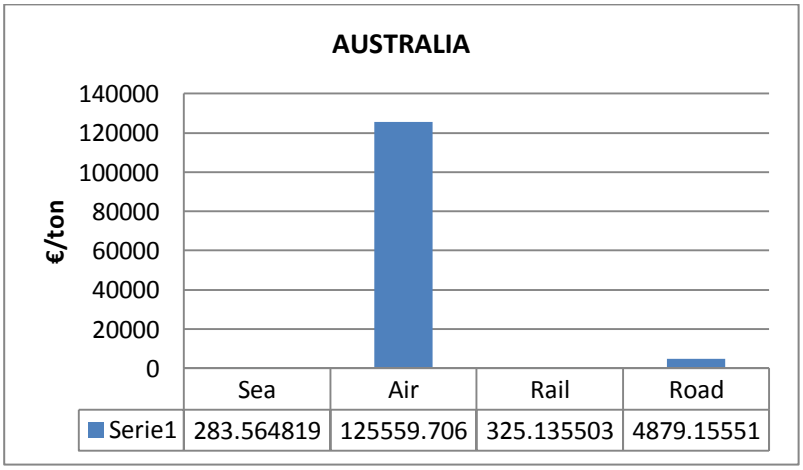

JAPAN
Sea
Air
Rail
Road
6%
5%
15%
74%

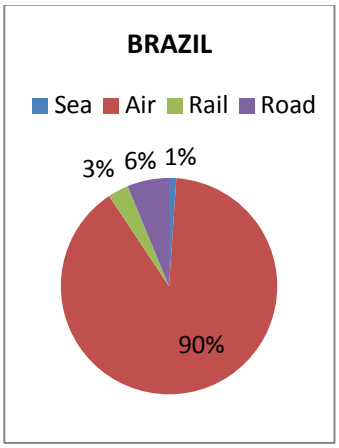

JAPAN
€/ton
120000
100000
80000
60000
40000
20000
0
Sea
Air
Rail
Road
Serie1
7727.48843
100885.552
7279.63589
20017.3743

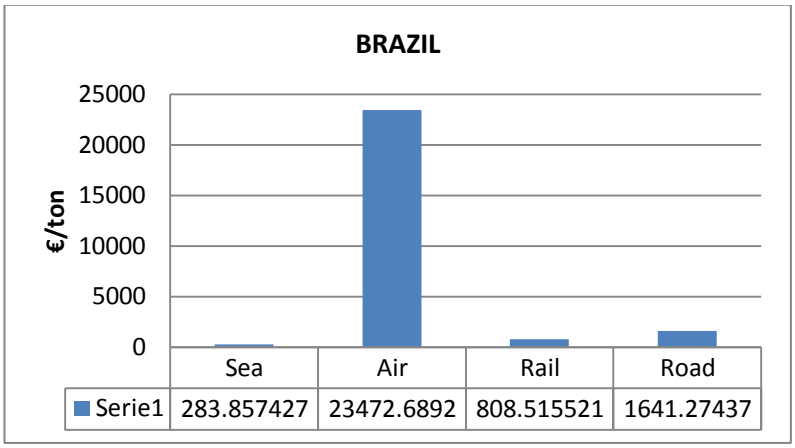

UNITED STATES
Sea
Air
Rail
Road
0%
1%
8%
91%

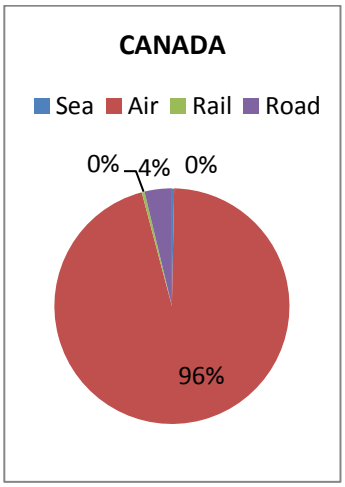

UNITED STATES
€/ton
140000
120000
100000
80000
60000
40000
20000
0
Sea
Air
Rail
Road
Serie1
864.184645
122717.511
465.566376
10491.8446

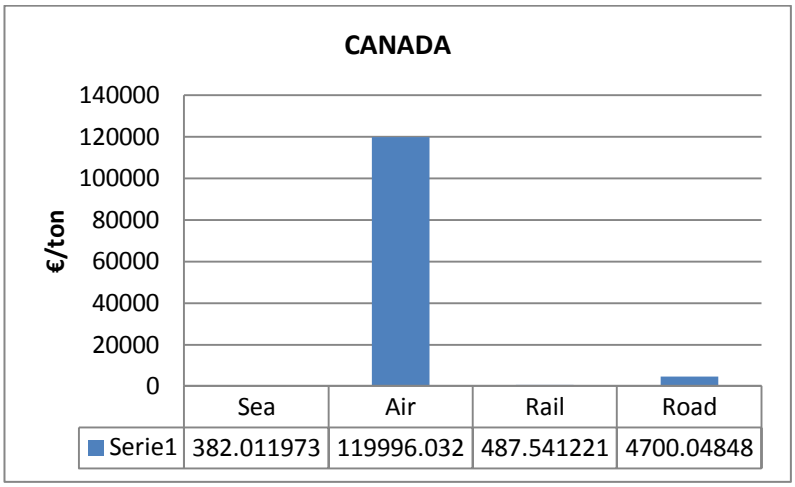

AUSTRALIA
Sea
Air
Rail
Road
4%
0%
0%
96%

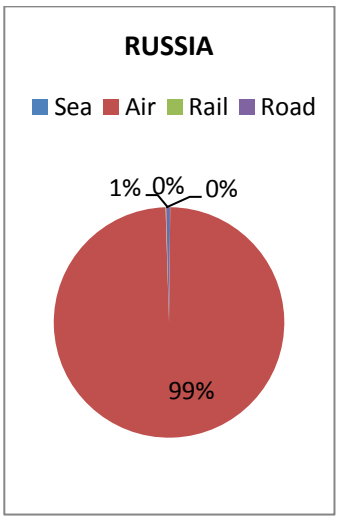

AUSTRALIA
€/ton
140000
120000
100000
80000
60000
40000
20000
0
Sea
Air
Rail
Road
Serie1
283.564819
125559.706
325.135503
4879.15551

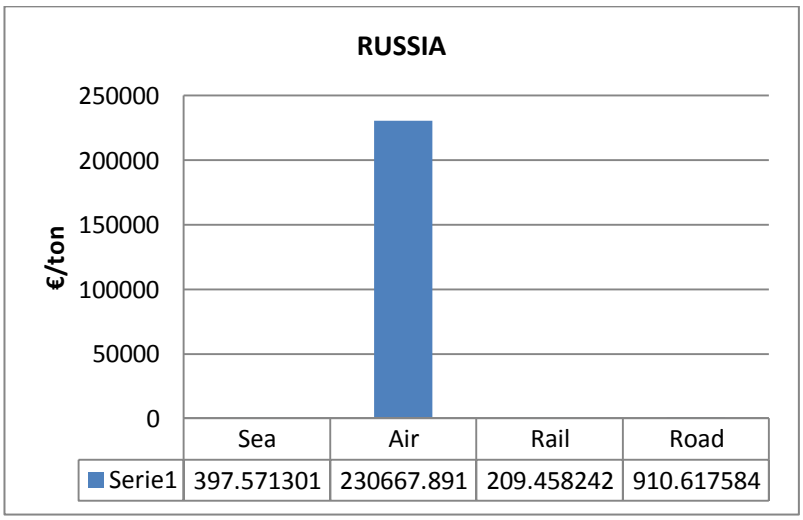

BRAZIL
Sea
Air
Rail
Road
3%
6%
1%
90%

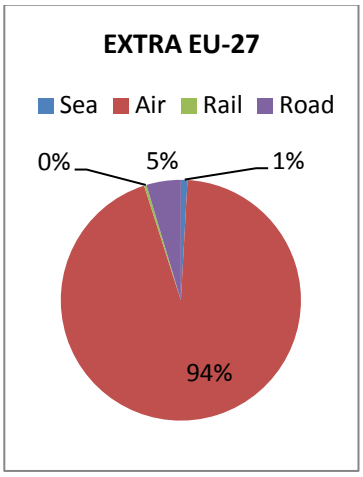

BRAZIL
€/ton
25000
20000
15000
10000
5000
0
Sea
Air
Rail
Road
Serie1
283.857427
23472.6892
808.515521
1641.27437

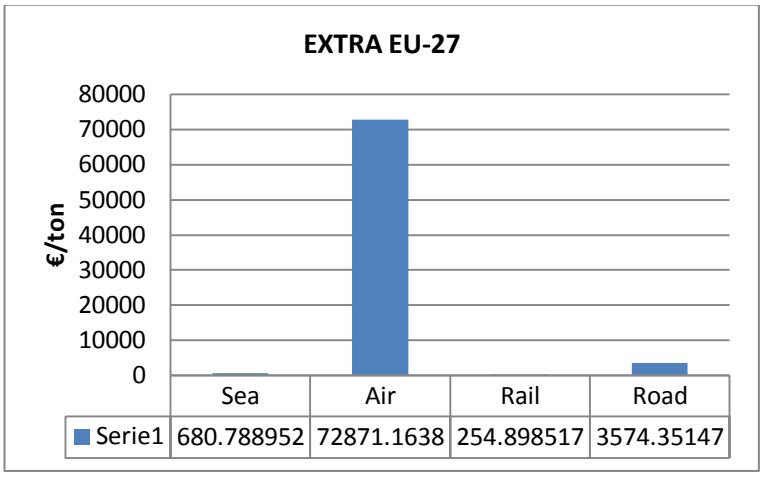

CANADA
Sea
Air
Rail
Road
0%
4%
0%
96%


CANADA
€/ton
140000
120000
100000
80000
60000
40000
20000
0
Sea
Air
Rail
Road
Serie1
382.011973
119996.032
487.541221
4700.04848


RUSSIA
Sea
Air
Rail
Road
1%
0%
0%
99%


RUSSIA
€/ton
250000
200000
150000
100000
50000
0
Sea
Air
Rail
Road
Serie1
397.571301
230667.891
209.458242
910.617584


EXTRA EU-27
Sea
Air
Rail
Road
0%
5%
1%
94%


EXTRA EU-27
€/ton
80000
70000
60000
50000
40000
30000
20000
10000
0
Sea
Air
Rail
Road
Serie1
680.788952
72871.1638
254.898517
3574.35147

APPENDIX C

Data: Variables Definition

PRODUCT CATEGORY	LIFE CYCLE	PERISHABILITY	TIMELINESS	TIME SENSITIVENESS	SPECIFIC WEIGHT (kg/dm^3) o (Kg/l)	BULKY	FRAGILITY&DANGEROUS	SATURATION LEVEL
001 - LIVE ANIMALS	1	3	1	1.8	1	1	1	0.5
011 - WHEAT, SPELT AND MESLIN	1	1	1	1	0.78	0	0	0.9
012 - BARLEY	1	1	1	1	0.65	0	0	0.9
013 - RYE	1	1	1	1	0.7	0	0	0.9
014 - OATS	1	1	1	1	0.64	0	0	0.9
015 - MAIZE	1	1	1	1	0.78	0	0	0.9
016 - RICE	1	1	1	1	0.7	0	0	0.9
019 - OTHER CEREALS N.E.S.	1	1	1	1	0.7	0	0	0.9
020 - POTATOES	1	2	1	1.4	0.7	0	0	0.9
031 - CITRUS FRUIT	1	3	1	1.8	n.a.	0	0	0.7
035 - OTHER FRUIT AND NUTS, FRESH	1	3	1	1.8	0.8	0	0	0.7
039 - OTHER VEGETABLES, FRESH OR FROZEN	1	3	1	1.8	n.a.	0	0	0.7
041 - WOOL AND OTHER ANIMAL HAIR	1	1	2	1.3	1.32	0	0	0.9

042 - COTTON	1	1	2	1.3	1.54	0	0	0.9
043 - MAN-MADE FIBRES	1	1	2	1.3	1.14	0	0	0.9
045 - SILK, FLAX, JUTE, TRUE HEMP AND OTHER VEGETABLE TEXTILE MATERIALS	1	1	2	1.3	1.33	0	0	0.9
049 - RAGS AND WASTE OF TEXTILE MATERIALS	1	1	2	1.3	1.3	0	0	0.9
051 - PAPER PULP WOOD	1	1	1	1	1.4	0	0	0.9
055 - OTHER WOOD IN THE ROUND	1	1	1	1	1.1	0	0	0.8
056 - RAILWAY OR TRAMWAY SLEEPERS OF WOOD AND OTHER WOOD ROUGHLY SQUARED, HALF SQUARED, OR SAWN	1	1	1	1	1	0	0	0.7
057 - FUEL WOOD, WOOD CHARCOAL, WOOD WASTE, CORK UNWORKED, WASTE CORK	1	1	1	1	1	0	0	0.8
091 - RAW HIDES AND SKINS, RAW FURSKINS, WASTE	1	1	2	1.3	n.a.	0	0	0.9
092 - RUBBER, NATURAL AND SYNTHETIC, RAW OR RECLAIMED	1	1	?	?	2	0	0	0.9
099 - OTHER NON-EDIBLE RAW VEGETABLE AND ANIMAL MATERIALS N.E.S.	1	3	?	?	n.a.	0	0	0.9
111 - RAW SUGAR	1	1	1	1	1.5	0	0	0.9
112 - REFINED SUGAR	1	1	1	1	1.6	0	0	0.9
113 - MOLASSES	1	1	1	1	1.4	0	0	0.9
121 - WINE OF FRESH GRAPES, GRAPE MUST	1	1	1	1	1.44	0	1	0.7
122 - BEER MADE FROM MALT	1	1	?	?	1.52	0	1	0.8
125 - OTHER ALCOHOLIC BEVERAGES	1	1	?	?	1.30	0	1	0.8
128 - NON-ALCOHOLIC BEVERAGES	1	1	?	?	1.44	0	1	0.8
131 - COFFEE	1	1	1	1	0.6	0	0	0.9
132 - COCOA AND CHOCOLATE	1	2	1	1.4	0.65	0	0	0.9
133 - TEA, MAT+, SPICES	1	1	1	1	0.43	0	0	0.9
134 - UNMANUFACTURED TOBACCO AND TOBACCO REFUSE	1	1	1	1	0.45	0	0	0.9

135 - MANUFACTURED TOBACCO	1	1	2	1.3	0.2	0	0	0.9
136 - GLUCOSE, DEXTROSE; OTHER SUGARS; SUGAR CONFECTIONERY; HONEY	1	2	1	1.4	1.4	0	0	0.9
139 - FOOD PREPARATIONS N.E.S.	1	3	2	2.1	n.a.	0	0	0.9
141 - MEAT, FRESH, CHILLED OR FROZEN	1	2	2	1.7	0.56	0	0	0.8
142 - FISH, CRUSTACEANS AND MOLLUSCS, FRESH, FROZEN, DRIED, SALTED OR SMOKED	1	3	2	2.1	0.6	0	0	0.8
143 - MILK AND CREAM, FRESH	1	3	3	2.4	0.56	0	0	0.9
144 - BUTTER, CHEESE, OTHER DAIRY PRODUCE	1	3	3	2.4	0.6	0	0	0.9
145 - MARGARINE, LARD AND EDIBLE FATS	1	3	3	2.4	0.92	0	0	0.9
146 - EGGS	1	3	3	2.4	0.37	0	0	0.7
147 - MEAT, DRIED, SALTED, SMOKED; PREPARED OR PRESERVED MEAT	1	2	2	1.7	0.56	0	0	0.8
148 - FISH, CRUSTACEANS AND MOLLUSCS, PREPARED OR PRESERVED	1	2	2	1.7	0.6	0	0	0.8
161 - FLOUR, CEREAL MEAL AND GROATS	1	1	2	1.3	0.7	0	0	0.9
162 - MALT	1	1	1	1	0.66	0	0	0.9
163 - OTHER CEREAL PREPARATIONS	1	1	2	1.3	0.69	0	0	0.9
164 - FRUIT, FROZEN, DRIED, DEHYDRATED; PREPARED AND PRESERVED FRUIT	1	1	2	1.3	n.a.	0	0	0.8
165 - DRIED VEGETABLES	1	1	1	1	n.a.	0	0	0.9
166 - PREPARED AND PRESERVED VEGETABLES	1	1	1	1	n.a.	0	0	0.9
167 - HOPS	1	1	1	1	0.66	0	0	0.9
171 - CEREAL STRAW, HAY AND HUSKS	1	1	1	1	0.6	0	0	0.9
172 - OIL-CAKE AND RESIDUES RESULTING FROM THE EXTRACTION OF VEGETABLE OILS	1	1	1	1	0.92	0	0	0.9
179 - BRAN, CEREAL BY-PRODUCTS AND OTHER ANIMAL FEEDINGSTUFFS N.E.S.; WASTE FROM FOOD INDUSTRIES	1	1	1	1	0.7	0	0	0.9
181 - OIL-SEED FATS, OILNUTS AND OIL KERNELS	1	1	1	1	n.a.	0	0	0.9

182 - ANIMAL AND VEGETABLE FATS AND OILS, AND PRODUCTS DERIVED THEREFROM	1	1	1	1	0.92	0	0	0.9
211 - COAL (ECSC)	1	1	1	1	0.7	0	0	0.9
213 - COAL BRIQUETTES (ECSC)	1	1	1	1	0.8	0	0	0.9
221 - LIGNITE (ECSC)	1	1	1	1	0.72	0	0	0.9
223 - LIGNITE BRIQUETTES (ECSC)	1	1	1	1	1	0	0	0.9
224 - PEAT	1	1	1	1	0.3	0	0	0.9
231 - COKE AND SEMI-COKE OF COAL (ECSC)	1	1	1	1	n.a.	0	0	0.9
233 - COKE AND SEMI-COKE OF LIGNITE (ECSC)	1	1	1	1	n.a.	0	1	0.9
310 - CRUDE PETROLEUM	1	1	1	1	0.8	0	1	0.9
321 - MOTOR SPIRIT	1	1	1	1	0.7	0	1	0.9
323 - KEROSENE, JET FUEL AND WHITE SPIRIT	1	1	1	1	0.76	0	1	0.9
325 - DISTILLATE FUELS	1	1	1	1	n.a.	0	1	0.9
327 - RESIDUAL FUEL OILS	1	1	1	1	n.a.	0	1	0.9
330 - GASEOUS HYDROCARBONS, LIQUID OR COMPRESSED	1	1	1	1	n.a.	0	0	0.9
341 - LUBRICATING OILS AND GREASES	1	2	1	1.4	0.92	0	0	0.9
343 - PETROLEUM BITUMEN AND BITUMINOUS MIXTURES	1	1	1	1	1.1	0	0	0.9
349 - OTHER NON-FUEL PETROLEUM DERIVATIVES	1	1	1	1	n.a.	0	0	0.9
410 - IRON ORE AND CONCENTRATES; EXCEPT ROASTED IRON PYRITES (ECSC)	1	1	1	1	2.56	0	0	0.8
451 - NON-FERROUS METAL WASTE	1	1	1	1	7	0	0	0.8
452 - COPPER ORE AND CONCENTRATES; COPPER MATTE	1	1	1	1	8	0	0	0.8
453 - BAUXITE AND CONCENTRATES	1	1	1	1	2.7	0	0	0.9
455 - MANGANESE ORE AND CONCENTRATES (ECSC)	1	1	1	1	7.21	0	0	0.9

459 - OTHER NON-FERROUS ORES AND CONCENTRATES	1	1	2	1.3	n.a.	0	0	0.9
462 - IRON AND STEEL WASTE FOR REMELTING (ECSC)	1	1	1	1	7.8	1	0	0.8
463 - IRON AND STEEL WASTE NOT FOR REMELTING (NON-ECSC)	1	1	1	1	7.8	0	0	0.8
465 - IRON SLAG FOR REMELTING (NON-ECSC)	1	1	1	1	7.8	1	0	0.8
466 - BLAST-FURNACE DUST (ECSC)	1	1	1	1	n.a.	0	0	0.9
467 - ROASTED IRON PYRITES (NON-ECSC)	1	1	1	1	n.a.	0	0	0.9
512 - PIG IRON, SPIEGELEISEN AND CARBURIZED FERRO-MANGANESE (ECSC)	1	1	1	1	7	0	0	0.9
513 - FERRO-ALLOYS OTHER THAN CARBURIZED FERRO-MANGANESE (NON-ECSC)	1	1	1	1	n.a.	0	0	0.9
515 - CRUDE STEEL (ECSC)	1	1	1	1	7.8	1	0	0.8
522 - SEMI-FINISHED ROLLED STEEL PRODUCTS (BLOOMS, BILLETS, SLABS, SHEET BARS, COILS) (ECSC)	1	1	1	1	7	1	0	0.7
523 - OTHER SEMI-FINISHED STEEL PRODUCTS (NON-ECSC)	1	1	1	1	7	1	0	0.7
532 - HOT-ROLLED OR -SHAPED STEEL (ECSC)	1	1	?	?	7	1	0	0.7
533 - COLD-ROLLED OR -SHAPED OR FORGED STEEL (NON-ECSC)	1	1	?	?	7	1	0	0.7
535 - WIRE ROD (ECSC)	1	1	?	?	7.6	1	0	0.8
536 - STEEL IRON AND STEEL WIRE (NON-ECSC)	1	1	?	?	7.5	1	0	0.8
537 - STEEL RAILS AND RAILWAY AND TRAMWAY TRACK-CONSTRUCTION MATERIAL (ECSC)	1	1	2	1.3	7.6	1	0	0.7
542 - SHEETS AND PLATES OF STEEL FOR RE-ROLLING; UNIVERSAL PLATES (ECSC)	1	1	?	?	7.5	1	0	0.7
543 - OTHER STEEL PLATES AND SHEETS (NON-ECSC)	1	1	?	?	7.5	1	0	0.7
545 - STEEL HOOP AND STRIP, TINPLATE (ECSC)	1	1	?	?	7.5	1	0	0.7
551 - TUBES, PIPES AND FITTINGS	1	1	?	?	n.a.	1	0	0.7
552 - IRON AND STEEL CASTINGS AND FORGINGS	1	1	?	?	n.a.	1	0	0.8
561 - COPPER AND COPPER ALLOYS, UNWROUGHT	1	1	?	?	8	1	0	0.8

562 - ALUMINIUM AND ALUMINIUM ALLOYS, UNWROUGHT	1	1	1	1	2.7	1	0	0.8
563 - LEAD AND LEAD ALLOYS, UNWROUGHT	1	1	1	1	11.3	1	0	0.8
564 - ZINC AND ZINC ALLOYS, UNWROUGHT	1	1	1	1	7.1	1	0	0.8
565 - OTHER NON-FERROUS METALS AND ALLOYS THEREOF, UNWROUGHT	1	1	?	?	n.a.	1	0	0.8
568 - FINISHED AND SEMI-FINISHED PRODUCTS OF NON-FERROUS METALS (EXCEPT MANUFACTURES)	1	1	2	1.3	n.a.	1	0	0.7
611 - SAND FOR INDUSTRIAL USE	1	1	1	1	1.6	0	0	0.9
612 - ORDINARY SAND AND GRAVEL	1	1	1	1	1.6	1	0	0.8
613 - PUMICE STONE, INCLUDING PUMICEOUS SAND AND GRAVEL	1	1	1	1	1.5	1	0	0.7
614 - CLAY AND CLAY EARTH	1	1	1	1	1.8	1	0	0.8
615 - SLAG NOT FOR RECOVERY OF METALS; ASH; DROSS	1	1	1	1	n.a.	0	0	0.8
621 - SALT, CRUDE OR REFINED	1	1	1	1	0.78	0	0	0.9
622 - UNROASTED IRON PYRITES	1	1	1	1	n.a.	0	0	0.9
623 - SULPHUR	1	1	1	1	2	0	0	0.9
631 - CRUSHED OR BROKEN STONE; PEBBLES, MACADAM, TARRED MACADAM	1	1	1	1	1.5	1	0	0.7
632 - BUILDING AND MONUMENTAL STONE, UNWORKED	1	1	1	1	2.7	1	0	0.7
633 - CALCAREOUS STONE FOR INDUSTRIAL PURPOSES	1	1	1	1	2.3	1	0	0.7
634 - CHALK	1	1	1	1	2.3	0	0	0.7
639 - OTHER CRUDE MINERALS	1	1	?	?	n.a.	0	0	0.8
641 - CEMENT	1	1	1	1	1.4	0	0	0.8
642 - LIME	1	1	1	1	1.8	0	0	0.8
650 - PLASTERS	1	?	1	?	2.3	0	0	0.8
691 - PUMICE STONE AGGLOMERATES; CONCRETE, CEMENT AND SIMILAR BUILDING MATERIALS	1	?	1	?	n.a.	1	0	0.7

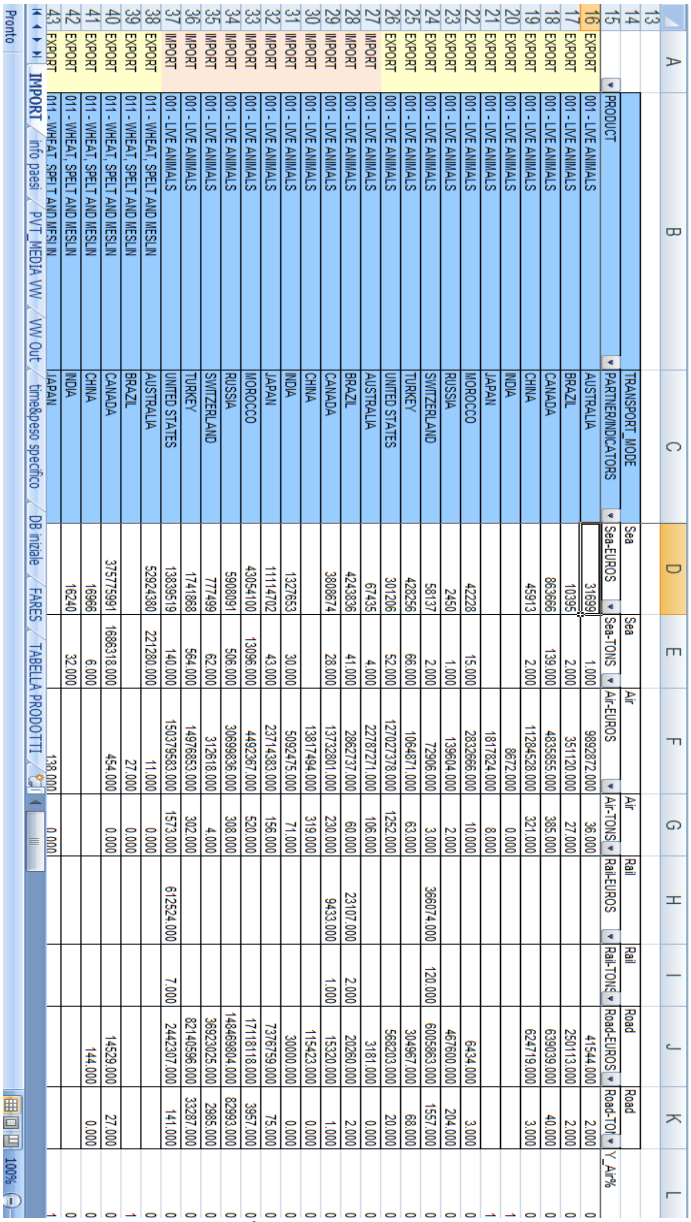
692 - BRICKS, ROOFING TILES AND OTHER CERAMIC BUILDING MATERIALS, REFRACTORY BUILDING MATERIALS	1	1	1	1	2.8	1	0	0.7
711 - SODIUM NITRATE, NATURAL	1	1	1	1	0.97	0	0	0.9
712 - PHOSPHATES, CRUDE, NATURAL	1	1	1	1	1.54	0	0	0.9
719 - OTHER NATURAL FERTILIZERS	1	1	1	1	n.a.	0	0	0.9
722 - OTHER PHOSPHATIC FERTILIZERS	1	1	1	1	n.a.	0	0	0.9
723 - POTASSIC FERTILIZERS	1	1	1	1	n.a.	0	0	0.9
724 - NITROGENOUS FERTILIZERS	1	1	1	1	n.a.	0	1	0.9
729 - COMPOSITE AND OTHER MANUFACTURED FERTILIZERS	1	1	1	1	n.a.	0	0	0.9
811 - SULPHURIC ACID; OLEUM	1	1	1	1	1.89	0	1	0.9
812 - CAUSTIC SODA AND SODA LYE	1	1	?	?	n.a.	0	1	0.9
813 - SODIUM CARBONATE (SODA ASH)	1	1	1	1	0.8	0	1	0.9
814 - CALCIUM CARBIDE	1	1	1	1	1.54	0	1	0.9
819 - OTHER BASIC CHEMICALS	1	1	3	1.6	n.a.	0	1	0.9
820 - ALUMINIUM OXIDE AND HYDROXIDE	1	1	?	?	n.a.	0	1	0.9
831 - BENZOLE	1	1	1	1	0.9	0	0	0.9
839 - PITCH, MINERAL TAR AND OTHER CRUDE MINERAL CHEMICAL DERIVATIVES FROM COAL AND NATURAL GAS	1	1	1	1	n.a.	0	0	0.9
841 - PAPER PULP	1	1	1	1	1.4	0	0	0.9
842 - WASTE PAPER AND SCRAP ARTICLES OF PAPER	1	1	1	1	0.7	0	0	0.9
891 - PLASTIC MATERIALS, UNWORKED	1	1	2	1.3	0.75	0	0	0.8
892 - DYEING, TANNING AND COLOURING MATERIALS	1	1	2	1.3	n.a.	0	1	0.9
893 - MEDICINAL AND PHARMACEUTICAL PRODUCTS; PERFUMERY AND CLEANSING PREPARATIONS	1	2	**3**	2	n.a.	0	1	0.8
894 - MANUFACTURED EXPLOSIVES, FIREWORKS AND OTHER PYROTECHNIC	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	1	0.7

ARTICLES, SPORTING AMMUNITION								
895 - STARCHES AND GLUTEN	1	1	2	1.3	0.7	0	1	0.9
896 - OTHER CHEMICAL PRODUCTS AND PREPARATIONS	1	2	3	2	n.a.	0	0	0.9
910 - TRANSPORT EQUIPMENT, WHETHER OR NOT ASSEMBLED, PARTS THEREOF	1	1	2	1.3	7	1	0	0.7
920 - TRACTORS; AGRICULTURAL MACHINERY AND EQUIPMENT, WHETHER OR NOT ASSEMBLED; PARTS THEREOF	1	1	2	1.3	7	1	0	0.7
931 - ELECTRICAL MACHINERY, APPARATUS, APPLIANCES AND ENGINES; PARTS THEREOF	3	1	3	2.2	n.a.	1	0	0.7
939 - NON-ELECTRICAL MACHINERY, APPARATUS, APPLIANCES AND ENGINES; PARTS THEREOF	2	1	2	1.6	n.a.	1	0	0.7
941 - FINISHED STRUCTURAL PARTS AND STRUCTURES	1	1	2	1.3	n.a.	1	0	0.7
949 - OTHER MANUFACTURES OF METAL	2	1	3	1.9	7		0	0.7
951 - GLASS	1	1	1	1	2.5	0	1	0.8
952 - GLASSWARE, POTTERY AND OTHER MANUFACTURES OF MINERALS	2	1	1	1.3	2.5	0	1	0.7
961 - LEATHER, MANUFACTURES OF LEATHER, OF RAW HIDE AND SKINS	2	1	3	1.9	1.2	0	0	0.8
962 - TEXTILE YARN, FABRICS, MADE-UP ARTICLES AND RELATED PRODUCTS	1	1	3	1.6	1.3	0	0	0.8
963 - TRAVEL GOODS, CLOTHING, KNITTED AND CROCHETED GOODS, FOOTWEAR	3	1	3	2.2	1.3	0	0	0.8
971 - SEMI-FINISHED PRODUCTS AND MANUFACTURED ARTICLES OF RUBBER	1	1	1	1	n.a.		0	0.8
972 - PAPER AND PAPERBOARD, UNWORKED	1	1	2	1.3	1	0	0	0.8
973 - PAPER AND PAPERBOARD MANUFACTURES	1	1	2	1.3	1	0	0	0.8
974 - PAPER MATTER	2	1	2	1.6	1	0	0	0.8
975 - FURNITURE, NEW	1	1	1	1	n.a.	1	0	0.7
976 - WOOD AND CORK MANUFACTURES, EXCLUDING FURNITURE	1	1	1	1	n.a.	1	0	0.7
979 - OTHER MANUFACTURED ARTICLES N.E.S.	3	1	3	2.2	n.a.	1	0	0.7
994 - GOLD, COINS, MEDALS	1	1	3	1.6	19	0	0	0.8

APPENDIX D

Data: Database Extract

Here we report an extract of the database we developed with all the variables that we created. In total we get 3400 observations.



	TRANSPORT_MODE	PRODUCT	PARTNERINDICATORS	Sea-EUROS	Sea-TONS	Air-EUROS	Air-TONS	Rail-EUROS	Rail-TONS	Road-EUROS	Road-TON	Y_A#%
16	EXPORT	001 - LIVE ANIMALS	AUSTRALIA	31693	1.000	9892872.000	36.000			41544.000	2.000	0
17	EXPORT	001 - LIVE ANIMALS	BRAZIL	10395	2.000	3511120.000	27.000			25011.000	2.000	0
18	EXPORT	001 - LIVE ANIMALS	CANADA	863666	139.000	4835855.000	385.000			639039.000	40.000	0
19	EXPORT	001 - LIVE ANIMALS	CHINA	45913	2.000	11284528.000	321.000			624719.000	3.000	0
20	EXPORT	001 - LIVE ANIMALS	INDIA			8672.000	0.000					1
21	EXPORT	001 - LIVE ANIMALS	JAPAN			1817824.000	8.000					1
22	EXPORT	001 - LIVE ANIMALS	MOROCCO	42228	15.000	2832666.000	10.000			6434.000	3.000	0
23	EXPORT	001 - LIVE ANIMALS	RUSSIA	2450	1.000	139604.000	2.000			467600.000	204.000	0
24	EXPORT	001 - LIVE ANIMALS	SWITZERLAND	58137	2.000	72906.000	3.000	366074.000	120.000	6005863.000	1557.000	0
25	EXPORT	001 - LIVE ANIMALS	TURKEY	428256	66.000	1064871.000	63.000			304967.000	68.000	0
26	EXPORT	001 - LIVE ANIMALS	UNITED STATES	301206	52.000	12702737.000	1252.000			568203.000	20.000	0
27	IMPORT	001 - LIVE ANIMALS	AUSTRALIA	67435	4.000	2278727.000	106.000			3181.000	0.000	0
28	IMPORT	001 - LIVE ANIMALS	BRAZIL	4243836	41.000	2862737.000	60.000	23107.000	2.000	20260.000	2.000	0
29	IMPORT	001 - LIVE ANIMALS	CANADA	3808674	28.000	13732801.000	230.000	9433.000	1.000	15320.000	1.000	0
30	IMPORT	001 - LIVE ANIMALS	CHINA			13817494.000	319.000			115423.000	0.000	0
31	IMPORT	001 - LIVE ANIMALS	INDIA	1327653	30.000	5092475.000	71.000			30000.000	0.000	0
32	IMPORT	001 - LIVE ANIMALS	JAPAN	11114702	43.000	2371438.000	156.000			73767.000	75.000	0
33	IMPORT	001 - LIVE ANIMALS	MOROCCO	43054100	13096.000	4492367.000	520.000			171181.000	3957.000	0
34	IMPORT	001 - LIVE ANIMALS	RUSSIA	5908091	506.000	30699836.000	308.000			1494698.000	82993.000	0
35	IMPORT	001 - LIVE ANIMALS	SWITZERLAND	777499	62.000	312618.000	4.000			368230.000	2985.000	0
36	IMPORT	001 - LIVE ANIMALS	TURKEY	1741868	564.000	14976853.000	392.000			82140596.000	33287.000	0
37	IMPORT	001 - LIVE ANIMALS	UNITED STATES	13839519	140.000	15037983.000	1573.000	612524.000	7.000	2442307.000	141.000	0
38	EXPORT	011 - WHEAT, SPELT AND MESLIN	AUSTRALIA	52924380	221280.000	11.000	0.000					0
39	EXPORT	011 - WHEAT, SPELT AND MESLIN	BRAZIL			27.000	0.000					1
40	EXPORT	011 - WHEAT, SPELT AND MESLIN	CANADA	375775991	1638318.000	454.000	0.000			14529.000	27.000	0
41	EXPORT	011 - WHEAT, SPELT AND MESLIN	CHINA	16986	6.000					144.000	0.000	0
42	EXPORT	011 - WHEAT, SPELT AND MESLIN	INDIA	16240	32.000							0
43	EXPORT	011 - WHEAT, SPELT AND MESLIN	JAPAN			138.000	0.000					1

Figure 51 - Data Base (1)

		PRODUCT	TRANSPORT_MODE	PARTNER/INDICATORS	Y_Air%_Out	Y%Probit	Value_Weigh	VW_Outliers	Distance	Connectivity_Airpl	Connectivity_Pave	GNI_per	Timeliness Dum	Timeline
					200	0.3							1.8	
16	EXPORT	001 - LIVE ANIMALS		AUSTRALIA	0.993	1	255541.410	255541.410	15935.000	0.656	0.387	38380.000	1.000	
17	EXPORT	001 - LIVE ANIMALS		BRAZIL	0.574	1	19729.935	19729.935	9395.881	1.292	0.055	10920.000	1.000	
18	EXPORT	001 - LIVE ANIMALS		CANADA	0.763	1	11238.582	11238.582	6541.714	0.901	0.400	37280.000	1.000	
19	EXPORT	001 - LIVE ANIMALS		CHINA	0.944	1	36672.270	36672.270	8031.667	3.134	0.540	7570.000	1.000	
20	EXPORT	001 - LIVE ANIMALS		INDIA	1.000	1	34688.000	34688.000	6565.789	1.825	0.490	3560.000	1.000	
21	EXPORT	001 - LIVE ANIMALS		JAPAN	1.000	1	227228.000	227228.000	9086.037	21.173	0.800	34780.000	1.000	
22	EXPORT	001 - LIVE ANIMALS		MOROCCO	0.983	1	102904.571	102904.571	2404.765	2.239	0.680	4620.000	1.000	
23	EXPORT	001 - LIVE ANIMALS		RUSSIA	0.229	0	2945.188	2945.188	2654.584	0.234	0.800	19190.000	1.000	
24	EXPORT	001 - LIVE ANIMALS		SWITZERLAND	0.011	0	3866.219	3866.219	543.470	48.444	1.000	48960.000	1.000	
25	EXPORT	001 - LIVE ANIMALS		TURKEY	0.592	1	9127.381	9127.381	2167.774	5.105	0.340	15180.000	1.000	
26	EXPORT	001 - LIVE ANIMALS		UNITED STATES	0.993	1	96598.782	96598.782	7595.452	9.602	0.670	47120.000	1.000	
27	IMPORT	001 - LIVE ANIMALS		AUSTRALIA	0.997	1	207798.973	207798.973	15935.090	0.656	0.387	38380.000	1.000	
28	IMPORT	001 - LIVE ANIMALS		BRAZIL	0.400	1	68094.667	68094.667	9395.881	1.292	0.055	10920.000	1.000	
29	IMPORT	001 - LIVE ANIMALS		CANADA	0.782	1	67562.415	67562.415	6541.714	0.901	0.400	37280.000	1.000	
30	IMPORT	001 - LIVE ANIMALS		CHINA	0.992	1	43676.856	43676.856	8031.667	3.134	0.540	7570.000	1.000	
31	IMPORT	001 - LIVE ANIMALS		INDIA	0.790	1	63862.653	63862.653	6565.789	1.825	0.490	3560.000	1.000	
32	IMPORT	001 - LIVE ANIMALS		JAPAN	0.562	1	154035.927	154035.927	9086.037	21.173	0.800	34780.000	1.000	
33	IMPORT	001 - LIVE ANIMALS		MOROCCO	0.069	0	3679.769	3679.769	2404.765	2.239	0.680	4620.000	1.000	
34	IMPORT	001 - LIVE ANIMALS		RUSSIA	0.166	0	2208.380	2208.380	2654.584	0.234	0.800	19190.000	1.000	
35	IMPORT	001 - LIVE ANIMALS		SWITZERLAND	0.008	0	12459.240	12459.240	543.470	48.444	1.000	48960.000	1.000	
36	IMPORT	001 - LIVE ANIMALS		TURKEY	0.151	0	2894.601	2894.601	2167.774	5.105	0.340	15180.000	1.000	
37	IMPORT	001 - LIVE ANIMALS		UNITED STATES	0.899	1	89883.897	89883.897	7595.452	9.602	0.670	47120.000	1.000	
38	EXPORT	011 - WHEAT, SPELT AND MESLIN		AUSTRALIA	0.000	0	239.174	239.174	15935.000	0.656	0.387	38380.000	0.000	
39	EXPORT	011 - WHEAT, SPELT AND MESLIN		BRAZIL			108.000	108.000	9395.881	1.292	0.055	10920.000	0.000	
40	EXPORT	011 - WHEAT, SPELT AND MESLIN		CANADA	0.000	0	222.843	222.843	6541.714	0.901	0.400	37280.000	0.000	
41	EXPORT	011 - WHEAT, SPELT AND MESLIN		CHINA	0.000	0	2851.667	2851.667	8031.667	3.134	0.540	7570.000	0.000	
42	EXPORT	011 - WHEAT, SPELT AND MESLIN		INDIA	0.000	0	507.500	507.500	6565.789	1.825	0.490	3560.000	0.000	
43	EXPORT	011 - WHEAT, SPELT AND MESLIN		JAPAN			552.000	552.000	9086.037	21.173	0.800	34780.000	0.000	

INPORT | info paesi | PVT_MEDIA VW | VW Out | time&peso specifico | DB iniziale | FARES | TABELLA PRODOTTI

Figure 52 - Data Base (2)

Data: Country Features

Country Name	Distance (km)	Connectivity: Airports Density	Connectivity: Paved Road (%of total road)	GDP (in PPP $)	GNI per capita in PPP in $	Time Difference [days]
AUSTRALIA	15935	0.0000006563	38.70%	865043	38380	32
BRAZIL	9395	0.0000012919	5.50%	2169180	10920	13.6
CANADA	6541	0.0000009014	40.00%	1327345	37280	7.2
CHINA	8031	0.0000031338	54.00%	10084764	7570	29
INDIA	6565	0.0000018252	49.00%	4198609	3560	16.75
JAPAN	9086	0.0000211733	80.00%	4332537	34780	31.4
MOROCCO	2404	0.0000022394	68.00%	151638	4620	3.09
RUSSIA	2703	0.0000002343	80.00%	2812383	19190	1.4
SWITZERLAND	689	0.0000484437	100.00%	361643	48960	-0.5
TURKEY	2167	0.0000051049	34.00%	1115994	15180	1.81
UNITED STATES	7595	0.0000096024	67.00%	14582400	47120	8.75

Data: Distance Correction

Here we report the spreadsheet for Switzerland case.

REPORTER/FLOW	Weight	Distance	Coefficient
AUSTRIA	0.07	576.41	39.37
BELGIUM	0.04	503.07	18.56
BULGARIA	0.00	1437.97	2.18
CYPRUS	0.00	2515.40	0.82
CZECH REPUBLIC	0.01	667.15	9.08
GERMANY	0.39	**543.47**	209.92
DENMARK	0.01	1015.16	7.16
ESTONIA	0.00	1789.41	0.86
SPAIN	0.04	1176.60	43.17
FINLAND	0.01	1922.12	11.08
FRANCE	0.12	473.73	56.11
UNITED KINGDOM	0.07	902.28	58.90
GREECE	0.01	1610.49	8.33
HUNGARY	0.01	874.42	5.60
IRELAND	0.02	1246.23	30.45
ITALY	0.14	609.68	87.71
LITHUANIA	0.00	1447.40	0.73
LUXEMBOURG	0.00	322.15	1.22
LATVIA	0.00	1568.27	1.32
MALTA	0.00	1362.43	0.90
NETHERLANDS	0.04	602.44	23.80
POLAND	0.01	1007.32	9.43
PORTUGAL	0.00	1633.79	6.41
ROMANIA	0.00	1379.99	4.91
SWEDEN	0.01	1422.90	15.88
SLOVENIA	0.00	553.12	1.18
SLOVAKIA	0.00	860.51	3.52
EU27	1.00		**658.59**

Data: Statistical Corrections

Data Base Structure: Observations and Dependent Variable

As already said in the thesis, our analysis started from the Eurostat database. The dependent variable "**Y%**" is the percentage of trade carried out by air. Any import or export voice in DB constitutes an observation in our analysis.

In our opinion there were observations that could bias the model. For instance the DB reported that in 2010 Europe imported 13 euros of petroleum from Australia and 50 euros from Brasil by air. For both of the observation this resulting in a unitary value of the dependent value (there is no import by any other mode of transport). These observations can bias the analysis because looking at the dependent variable it seems very likely for petroleum to be delivered by air from Australia and Brazil. But is this result actually reliable?

Those observations referring to very small import/export values could be the result of gathering mistakes. And even if they are not mistakes, it's dangerous to equally weight them in the regression analysis (there could be many different and unexpected reasons why these data could show up).

Hence we created the new adjusted dependent variable "**Y%outliers**" by setting arbitrary bounds referring to minimum amount of import/export to detect automatically outliers in Excel.

Since Switzerland and Russia are neighbouring countries for the European Union, we decided to create another adjusted independent variable labelled "**Yspeed%**"; this variable indicates the probability of selecting the fastest transport mode that in case of Switzerland and Russia could be both planes and trucks.

All these dependent variables will be used and tested in the next correlation and regression analyses.

The "Value/Weight Problem"

As already said above, the ratio between value (in Euros) and weight (in tons) of trade is a very important indicator for the analysis.

Software is not able to compute the value/weight for some observation reporting "DIV/#0". It happens when observations have trade weight equal to "0", but a positive value in Euros, so that it's impossible to compute the ratio since the denominator is null.

The problem is that if a product is traded, Eurostat reports any positive value in Euros (even very low values), while they report a null value if import/export weights are lower than half a ton.

To enter data in the statistic software we need to get rid of the DIV/#0 signs, so that a first option can be to just eliminate these observations considering them outliers. The problem with this approach is that we are strongly reducing the size of the sample.

Secondly we thought to substitute the null value with a parameter, so to compute an "arbitrary value/weight". The problem with this second approach is that if the parameter isn't a good estimation (the actual data can range from 0 to 0,5 tons in absence of mistakes) we are distorting data.

A third solution is to enter the parameter and then to look for statistical evidences about the likelihood of the "hypotheses".

To implement this third approach, we computed V/W mean and standard deviation for any product categories. Then we computed the standard error (good estimator for the "sampling standard deviation"); finally we built

Confidence Intervals referring to “Student distribution”. We can accept only those “arbitrary value-weight” that are within the interval.

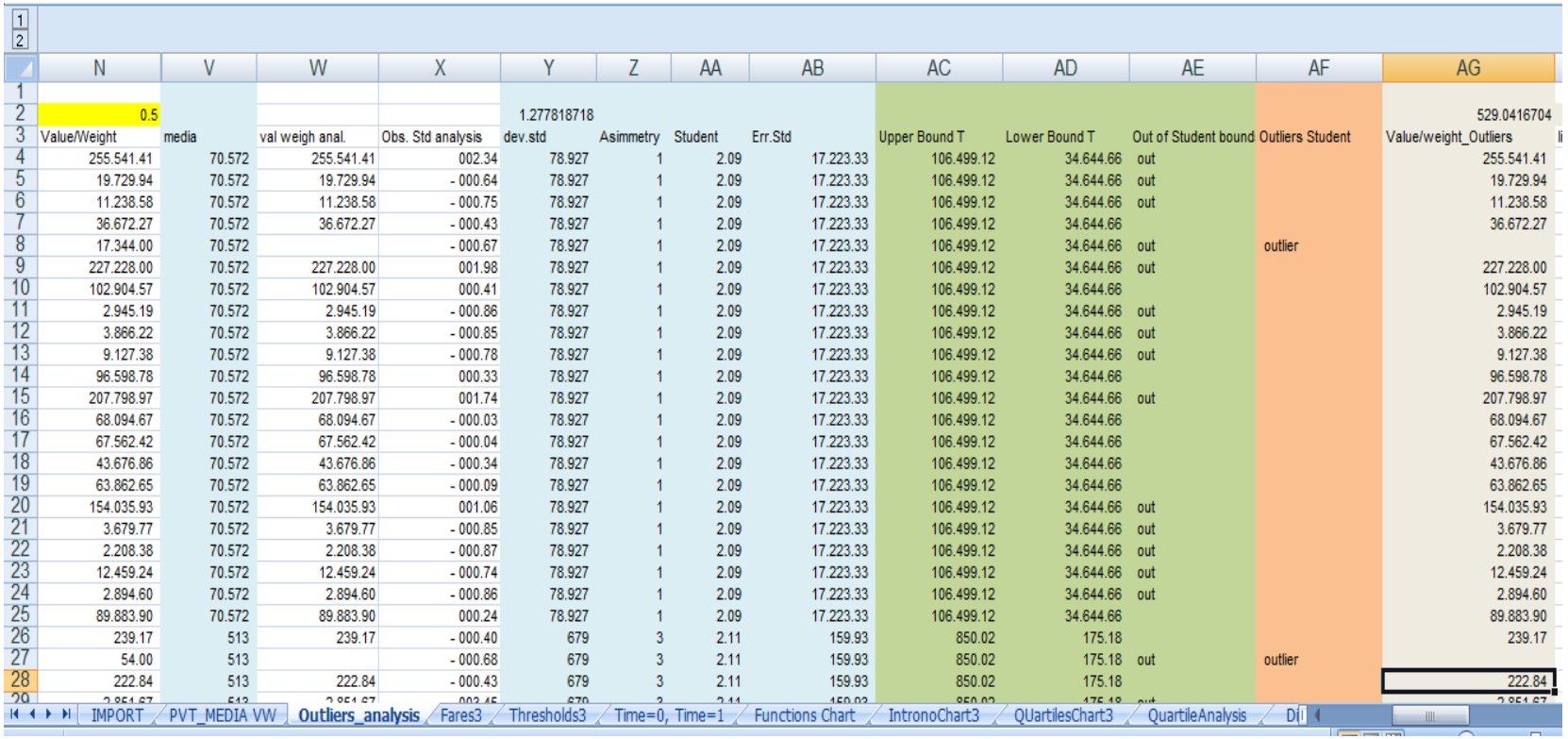
Confidence Interval → $X \pm t\,(1 - \alpha/2;\ n\text{-}1) * \delta/\sqrt{n}$

We started considering a 0.25 (Tons) arbitrary weight, and we built the acceptance intervals with a 0.95 confidence. However, we can automatically re-compute outliers while changing parameters in the Excel spreadsheet.

The Student distribution looks to be adept at our situation: since the indicator to be tested is already a statistical value, we can assume that the population is distributed as a normal distribution. Moreover, for small samples (we can have a maximum number of observation equal to 22, 11 import values plus 11 export values) the Student distribution is more precise than the standard one.

However, we computed “distribution asymmetries” to check if distributions can be assumed as normal, finding that asymmetry is quite low: it always varies from 0 to 4 in our samples.

In the following picture we show the Excel spreadsheets that we develop for “Value/Weight” statistical corrections:



	N	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG
1													
2	0.5				1.277818718								529.0416704
3	Value/Weight	media	val weigh anal.	Obs. Std analysis	dev.std	Asimmetry	Student	Err.Std	Upper Bound T	Lower Bound T	Out of Student bound	Outliers Student	Value/weight_Outliers
4	255.541.41	70.572	255.541.41	002.34	78.927	1	2.09	17.223.33	106.499.12	34.644.66	out		255.541.41
5	19.729.94	70.572	19.729.94	- 000.64	78.927	1	2.09	17.223.33	106.499.12	34.644.66	out		19.729.94
6	11.238.58	70.572	11.238.58	- 000.75	78.927	1	2.09	17.223.33	106.499.12	34.644.66	out		11.238.58
7	36.672.27	70.572	36.672.27	- 000.43	78.927	1	2.09	17.223.33	106.499.12	34.644.66			36.672.27
8	17.344.00	70.572		- 000.67	78.927	1	2.09	17.223.33	106.499.12	34.644.66	out	outlier	
9	227.228.00	70.572	227.228.00	001.98	78.927	1	2.09	17.223.33	106.499.12	34.644.66	out		227.228.00
10	102.904.57	70.572	102.904.57	000.41	78.927	1	2.09	17.223.33	106.499.12	34.644.66			102.904.57
11	2.945.19	70.572	2.945.19	- 000.86	78.927	1	2.09	17.223.33	106.499.12	34.644.66	out		2.945.19
12	3.866.22	70.572	3.866.22	- 000.85	78.927	1	2.09	17.223.33	106.499.12	34.644.66	out		3.866.22
13	9.127.38	70.572	9.127.38	- 000.78	78.927	1	2.09	17.223.33	106.499.12	34.644.66	out		9.127.38
14	96.598.78	70.572	96.598.78	000.33	78.927	1	2.09	17.223.33	106.499.12	34.644.66			96.598.78
15	207.798.97	70.572	207.798.97	001.74	78.927	1	2.09	17.223.33	106.499.12	34.644.66	out		207.798.97
16	68.094.67	70.572	68.094.67	- 000.03	78.927	1	2.09	17.223.33	106.499.12	34.644.66			68.094.67
17	67.562.42	70.572	67.562.42	- 000.04	78.927	1	2.09	17.223.33	106.499.12	34.644.66			67.562.42
18	43.676.86	70.572	43.676.86	- 000.34	78.927	1	2.09	17.223.33	106.499.12	34.644.66			43.676.86
19	63.862.65	70.572	63.862.65	- 000.09	78.927	1	2.09	17.223.33	106.499.12	34.644.66			63.862.65
20	154.035.93	70.572	154.035.93	001.06	78.927	1	2.09	17.223.33	106.499.12	34.644.66	out		154.035.93
21	3.679.77	70.572	3.679.77	- 000.85	78.927	1	2.09	17.223.33	106.499.12	34.644.66	out		3.679.77
22	2.208.38	70.572	2.208.38	- 000.87	78.927	1	2.09	17.223.33	106.499.12	34.644.66	out		2.208.38
23	12.459.24	70.572	12.459.24	- 000.74	78.927	1	2.09	17.223.33	106.499.12	34.644.66	out		12.459.24
24	2.894.60	70.572	2.894.60	- 000.86	78.927	1	2.09	17.223.33	106.499.12	34.644.66	out		2.894.60
25	89.883.90	70.572	89.883.90	000.24	78.927	1	2.09	17.223.33	106.499.12	34.644.66			89.883.90
26	239.17	513	239.17	- 000.40	679	3	2.11	159.93	850.02	175.18			239.17
27	54.00	513		- 000.68	679	3	2.11	159.93	850.02	175.18	out	outlier	
28	222.84	513	222.84	- 000.43	679	3	2.11	159.93	850.02	175.18			222.84
29	2.851.67	513	2.851.67	003.45	679	3	2.11	159.93	850.02	175.18	out		2.851.67

IMPORT | PVT_MEDIA VW | Outliers_analysis | Fares3 | Thresholds3 | Time=0, Time=1 | Functions Chart | IntronoChart3 | QUartilesChart3 | QuartileAnalysis | D(

APPENDIX E

Data: "Fare Difference 3" Definition

Here we report the data about transportation costs that we gathered for "Fare Difference 3" definition.

SEA FREIGHT COSTS

We obtained these data from Artsana's interview:

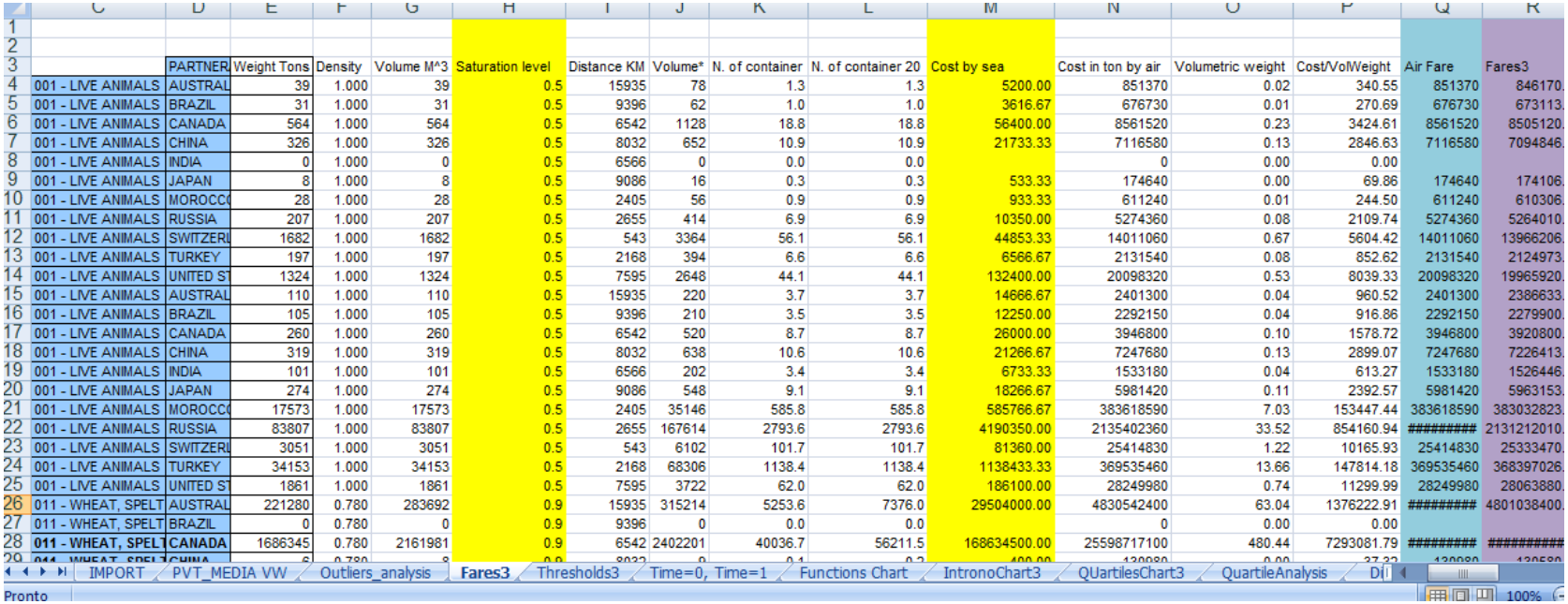
Destination	Distance [km]	Cost for one container [€]
AUSTRALIA	15935	4000
BRAZIL	9395.881	3500
CANADA	6541.714	3000
CHINA	8031.667	2000
INDIA	6565.789	2000
JAPAN	9086.037	2000
MOROCCO	2404.765	1000
RUSSIA	2654.584	1500
SWITZERLAND	543.47	800
TURKEY	2167.774	1000
UNITED STATES	7595.452	3000

AIR SHIPPING COSTS

In the following table we report the "air shipping costs" related to the UPS rate table. Fares change between European import and export.

Destination	Export cost/kg by UPS	Import cost/Kg by UPS	EXP Cost/ton by air	IMP Cost/ton by air
AUSTRALIA	21.83	21.83	21830	21830
BRAZIL	21.83	21.83	21830	21830
CANADA	15.18	15.18	15180	15180
CHINA	21.83	22.72	21830	22720
INDIA	21.83	15.18	21830	15180
JAPAN	21.83	21.83	21830	21830
MOROCCO	24.75	21.83	24750	21830
RUSSIA	25.48	25.48	25480	25480
SWITZERLAND	8.33	8.33	8330	8330
TURKEY	10.82	10.82	10820	10820
UNITED STATES	15.18	15.18	15180	15180

Here, we report a picture of the Excel spreadsheet that we created for "Fares Difference 3" computations:



		PARTNER	Weight Tons	Density	Volume M^3	Saturation level	Distance KM	Volume*	N. of container	N. of container 20	Cost by sea	Cost in ton by air	Volumetric weight	Cost/VolWeight	Air Fare	Fares3
4	001 - LIVE ANIMALS	AUSTRAL	39	1.000	39	0.5	15935	78	1.3	1.3	5200.00	851370	0.02	340.55	851370	846170.
5	001 - LIVE ANIMALS	BRAZIL	31	1.000	31	0.5	9396	62	1.0	1.0	3616.67	676730	0.01	270.69	676730	673113.
6	001 - LIVE ANIMALS	CANADA	564	1.000	564	0.5	6542	1128	18.8	18.8	56400.00	8561520	0.23	3424.61	8561520	8505120.
7	001 - LIVE ANIMALS	CHINA	326	1.000	326	0.5	8032	652	10.9	10.9	21733.33	7116580	0.13	2846.63	7116580	7094846.
8	001 - LIVE ANIMALS	INDIA	0	1.000	0	0.5	6566	0	0.0	0.0		0	0.00	0.00		
9	001 - LIVE ANIMALS	JAPAN	8	1.000	8	0.5	9086	16	0.3	0.3	533.33	174640	0.00	69.86	174640	174106.
10	001 - LIVE ANIMALS	MOROCC	28	1.000	28	0.5	2405	56	0.9	0.9	933.33	611240	0.01	244.50	611240	610306.
11	001 - LIVE ANIMALS	RUSSIA	207	1.000	207	0.5	2655	414	6.9	6.9	10350.00	5274360	0.08	2109.74	5274360	5264010.
12	001 - LIVE ANIMALS	SWITZERL	1682	1.000	1682	0.5	543	3364	56.1	56.1	44853.33	14011060	0.67	5604.42	14011060	13966206.
13	001 - LIVE ANIMALS	TURKEY	197	1.000	197	0.5	2168	394	6.6	6.6	6566.67	2131540	0.08	852.62	2131540	2124973.
14	001 - LIVE ANIMALS	UNITED ST	1324	1.000	1324	0.5	7595	2648	44.1	44.1	132400.00	20098320	0.53	8039.33	20098320	19965920.
15	001 - LIVE ANIMALS	AUSTRAL	110	1.000	110	0.5	15935	220	3.7	3.7	14666.67	2401300	0.04	960.52	2401300	2386633.
16	001 - LIVE ANIMALS	BRAZIL	105	1.000	105	0.5	9396	210	3.5	3.5	12250.00	2292150	0.04	916.86	2292150	2279900.
17	001 - LIVE ANIMALS	CANADA	260	1.000	260	0.5	6542	520	8.7	8.7	26000.00	3946800	0.10	1578.72	3946800	3920800.
18	001 - LIVE ANIMALS	CHINA	319	1.000	319	0.5	8032	638	10.6	10.6	21266.67	7247680	0.13	2899.07	7247680	7226413.
19	001 - LIVE ANIMALS	INDIA	101	1.000	101	0.5	6566	202	3.4	3.4	6733.33	1533180	0.04	613.27	1533180	1526446.
20	001 - LIVE ANIMALS	JAPAN	274	1.000	274	0.5	9086	548	9.1	9.1	18266.67	5981420	0.11	2392.57	5981420	5963153.
21	001 - LIVE ANIMALS	MOROCC	17573	1.000	17573	0.5	2405	35146	585.8	585.8	585766.67	383618590	7.03	153447.44	383618590	383032823.
22	001 - LIVE ANIMALS	RUSSIA	83807	1.000	83807	0.5	2655	167614	2793.6	2793.6	4190350.00	2135402360	33.52	854160.94	#########	2131212010.
23	001 - LIVE ANIMALS	SWITZERL	3051	1.000	3051	0.5	543	6102	101.7	101.7	81360.00	25414830	1.22	10165.93	25414830	25333470.
24	001 - LIVE ANIMALS	TURKEY	34153	1.000	34153	0.5	2168	68306	1138.4	1138.4	1138433.33	369535460	13.66	147814.18	369535460	368397026.
25	001 - LIVE ANIMALS	UNITED ST	1861	1.000	1861	0.5	7595	3722	62.0	62.0	186100.00	28249980	0.74	11299.99	28249980	28063880.
26	011 - WHEAT, SPELT	AUSTRAL	221280	0.780	283692	0.9	15935	315214	5253.6	7376.0	29504000.00	4830542400	63.04	1376222.91	#########	4801038400.
27	011 - WHEAT, SPELT	BRAZIL	0	0.780	0	0.9	9396	0	0.0	0.0		0	0.00	0.00		
28	011 - WHEAT, SPELT	CANADA	1686345	0.780	2161981	0.9	6542	2402201	40036.7	56211.5	168634500.00	25598717100	480.44	7293081.79	#########	#########
29	011 - WHEAT, SPELT	CHINA	6	0.780	8	0.9	8032	0	0.1	0.2	400.00	130080	0.00	37.33	130080	129680.

IMPORT | PVT_MEDIA VW | Outliers_analysis | Fares3 | Thresholds3 | Time=0, Time=1 | Functions Chart | IntronoChart3 | QUartilesChart3 | QuartileAnalysis | D(

Pronto

100%

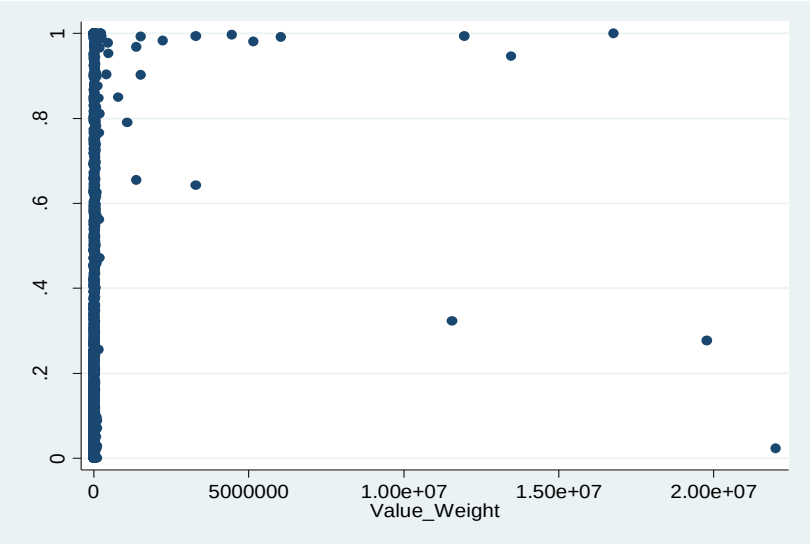
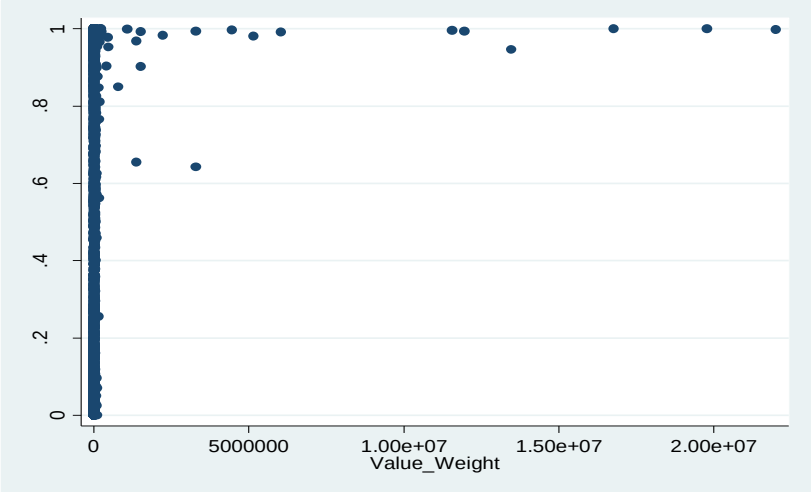
APPENDIX F

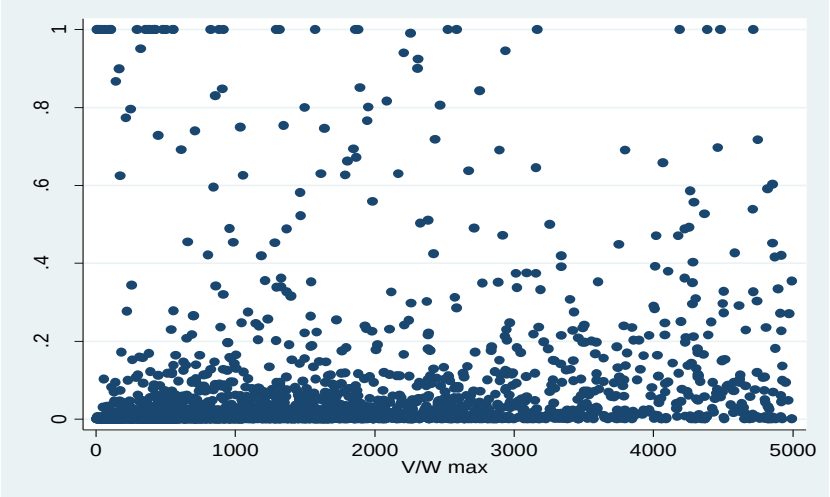
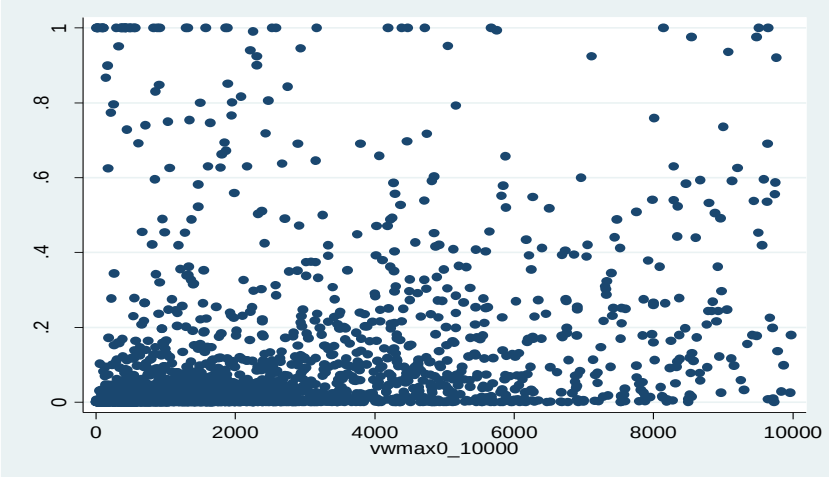
Correlation Analysis

We report the all the graphs about the correlation analysis that we implemented with STATA.

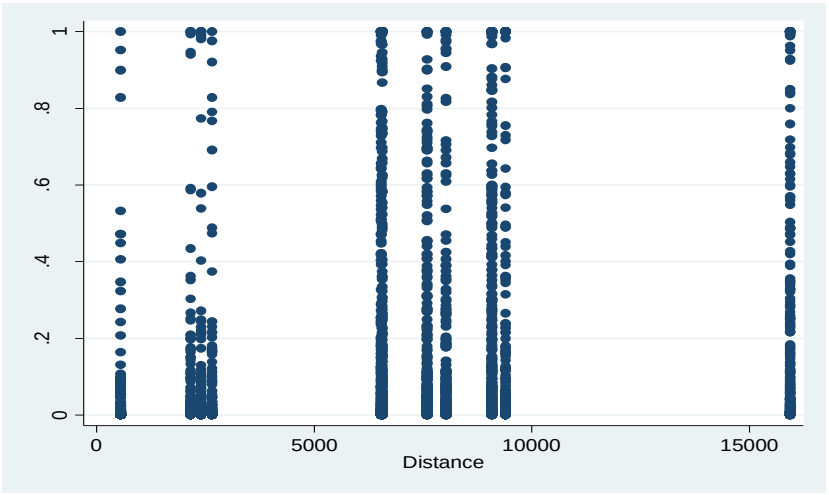
Correlation between the Dependent Variable (Y%) and all the other Independent Variables:

VALUE/WEIGHT

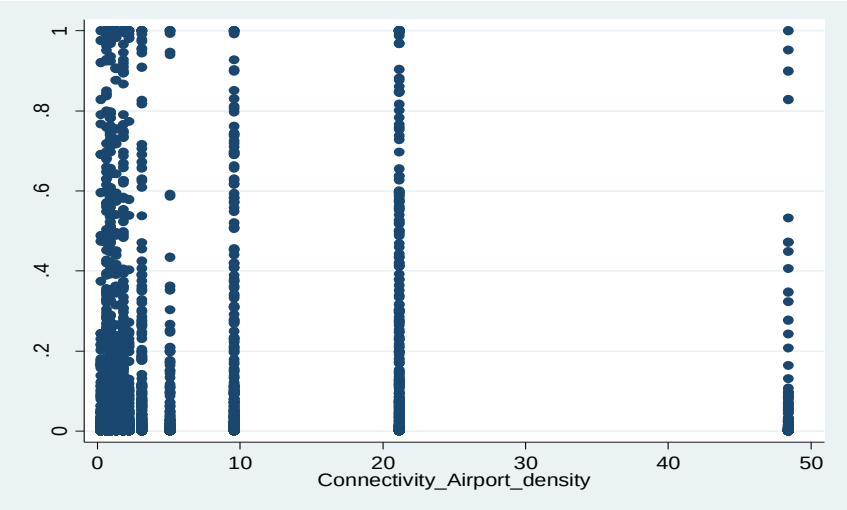
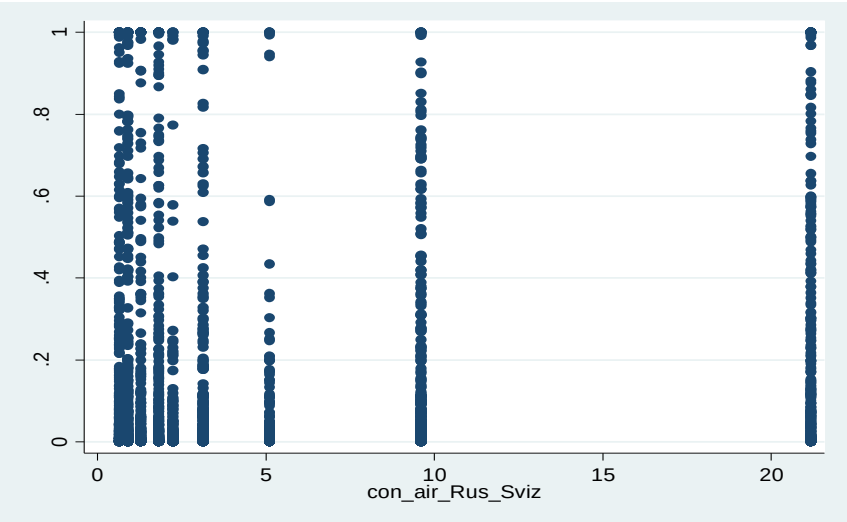
Y% → Value-Weight	0.119	
Yspeed% → Value/Weight	0.1144	

Y% → V/W-max (0-5000) We included just the observation with a V/W beteen 0 and 5000.	0.1851	
Y% → V/W-max (0-10000) We included just the observation with a V/W beteen 0 and 10000.	0.3	

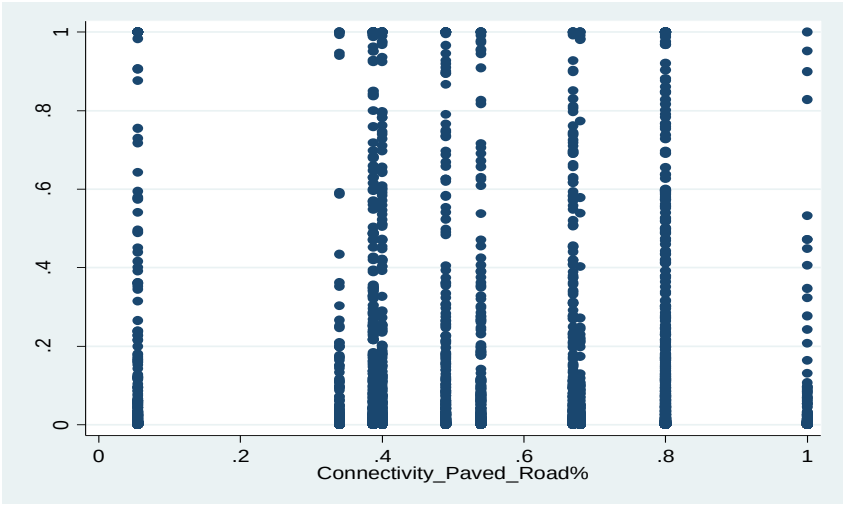
DISTANCE

Y% → Distance	0.2265	

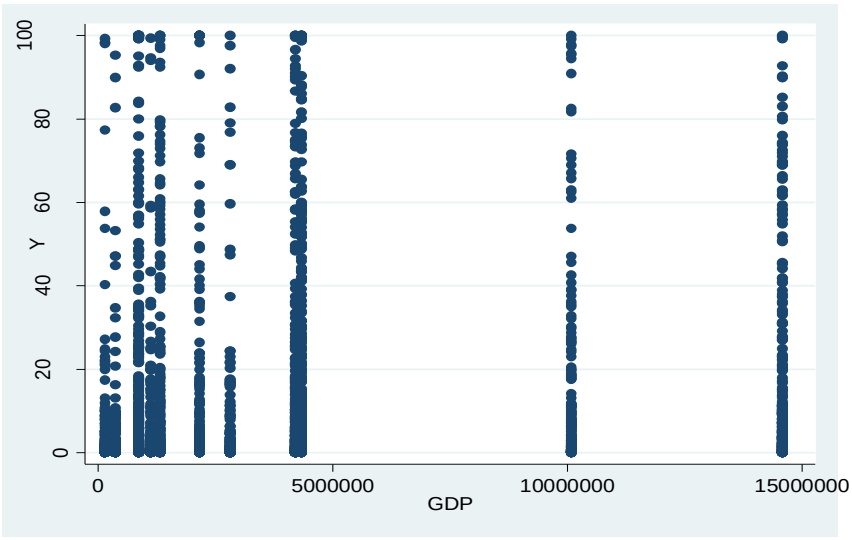
AIRPORT DENSITY

Y% →Airport Density	-0.056	
Y% → Airport density Russia and Switzerland are excluded	0.101	

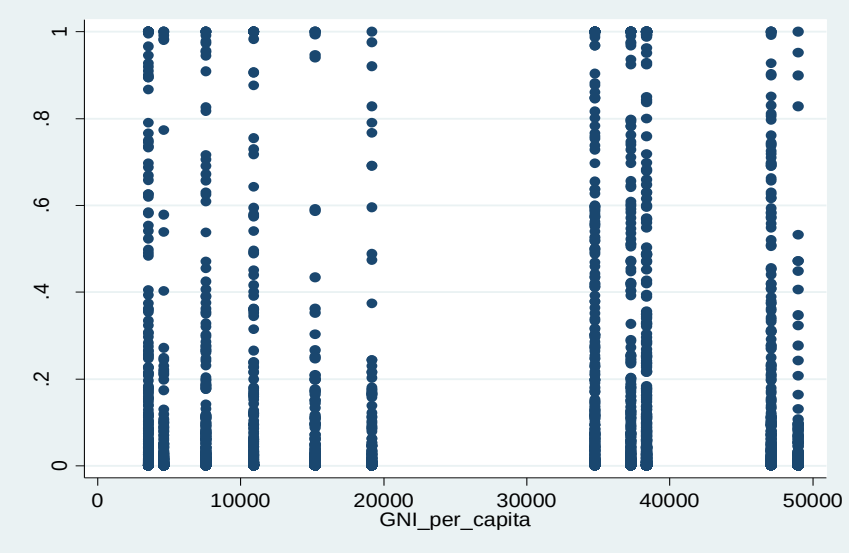
PAVED ROAD

Y% → Paved Road	-0.0754	

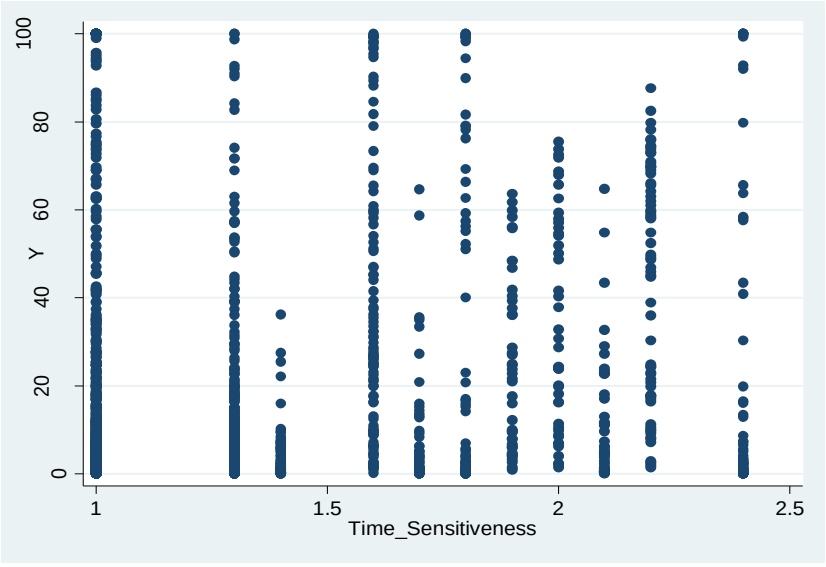
GDP

Y% → GDP	0.0749	

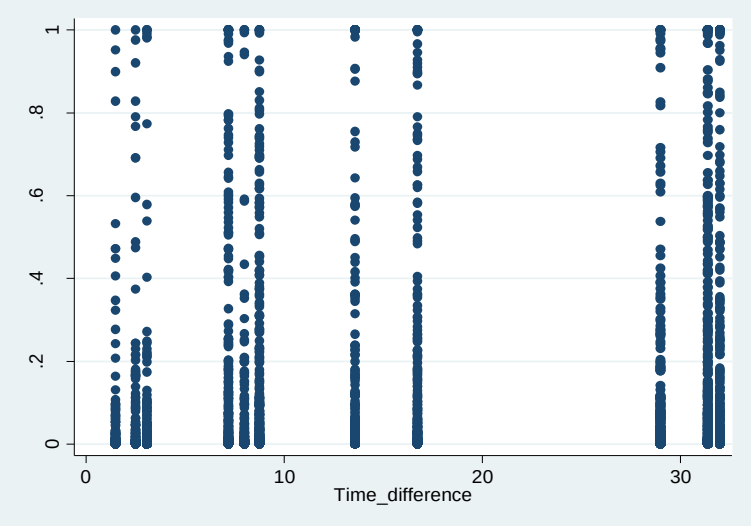
GNI/Capita

Y% → GNI/Capita	0.0664	

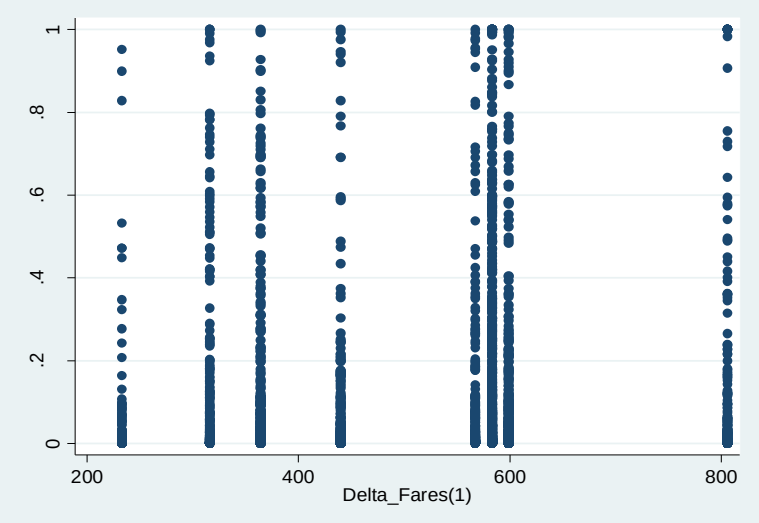
TIME SENSITIVENESS

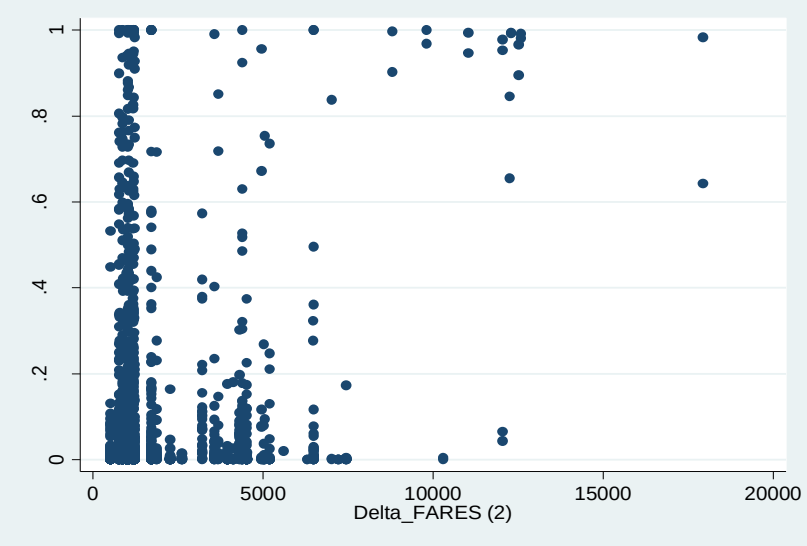
Y% → Time Sensitiveness	0.2484	

TIME DIFFERENCE

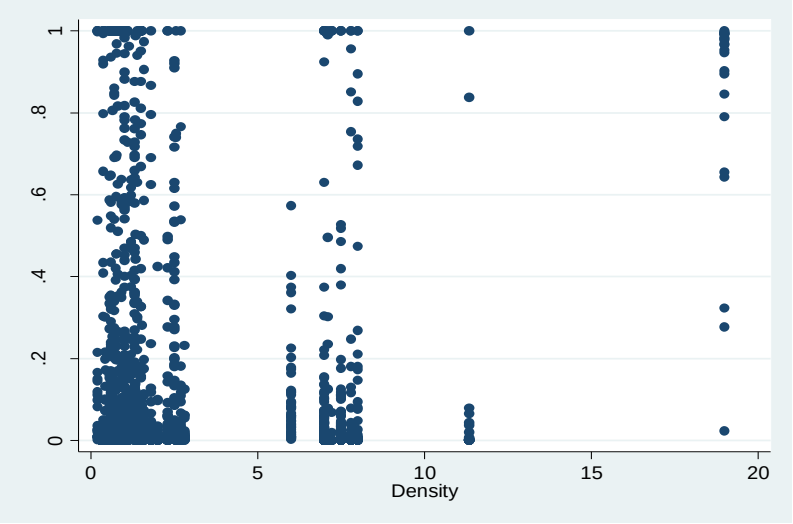
Y% → Time Difference	0.1988	

DELTA FARE

Y% → Delta Fare (1)	0.0942	

Y% → Delta Fare (2)	0.1758	

DENSITY

Y% → Density	0.1147	

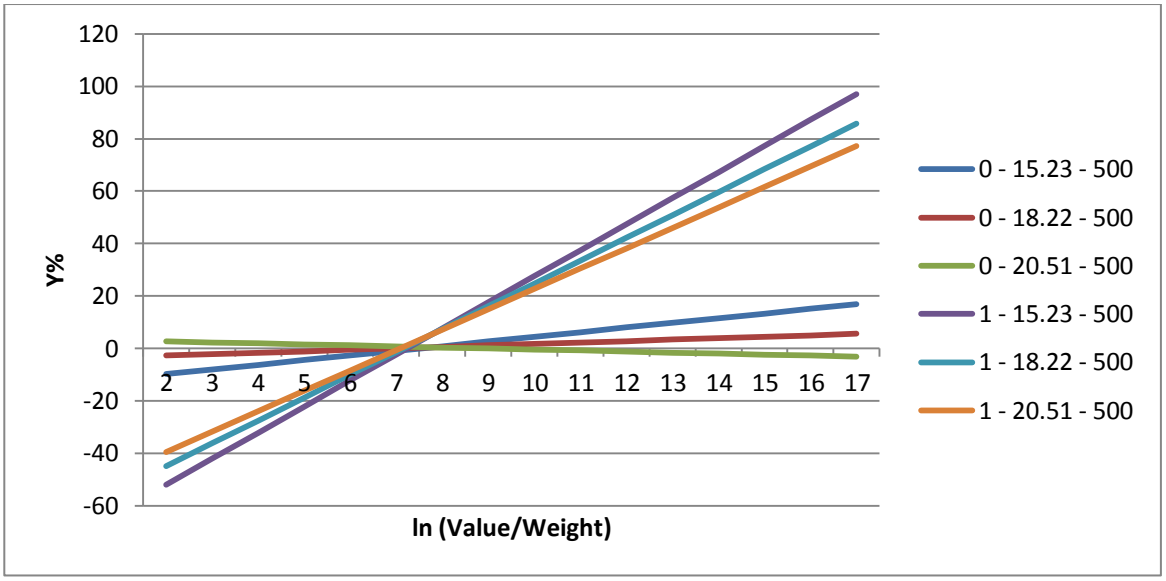
APPENDIX G

Variables Interdependences: The combined effect of Distance in Framework

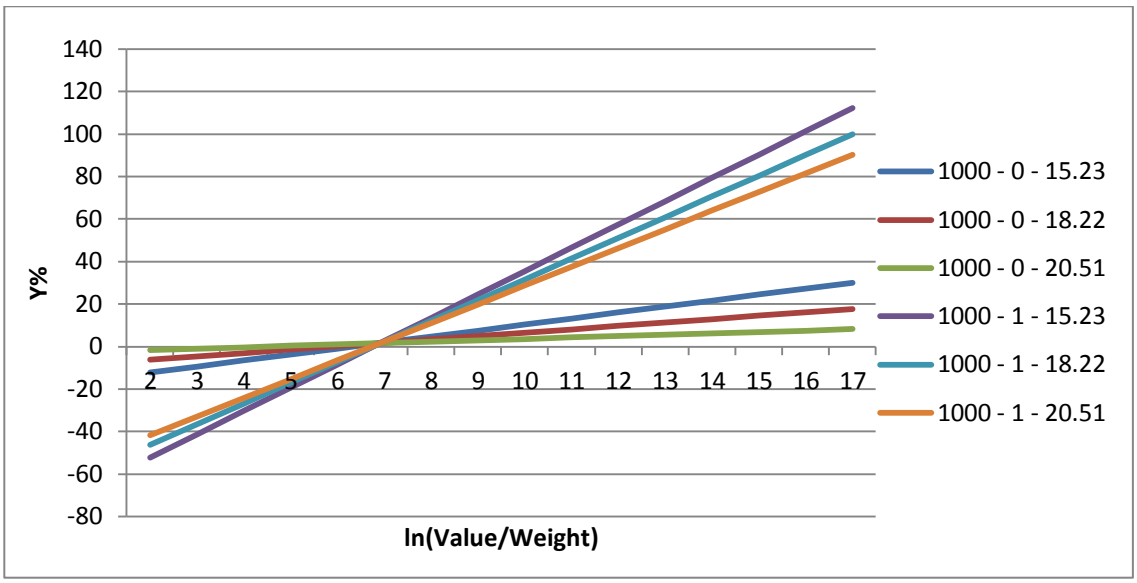
Here we report all of the graphs we made for the "Distance Analysis".

Any different line refers to different fare levels distinguishing between timely or not timely trade. In order to get a clearer view, we develop different graphs corresponding to different reference distances.

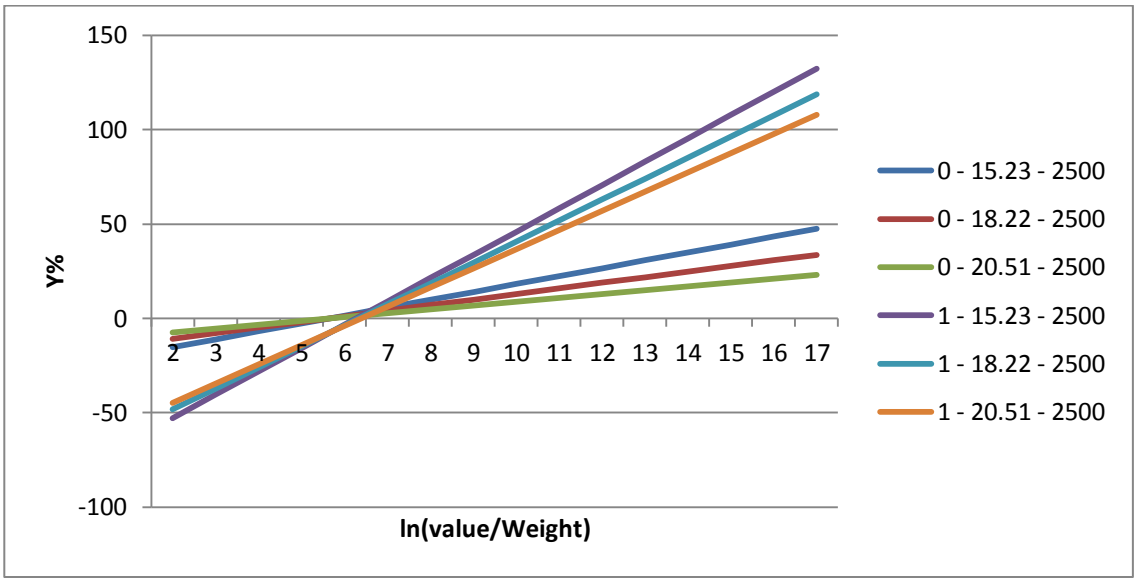
500 Km



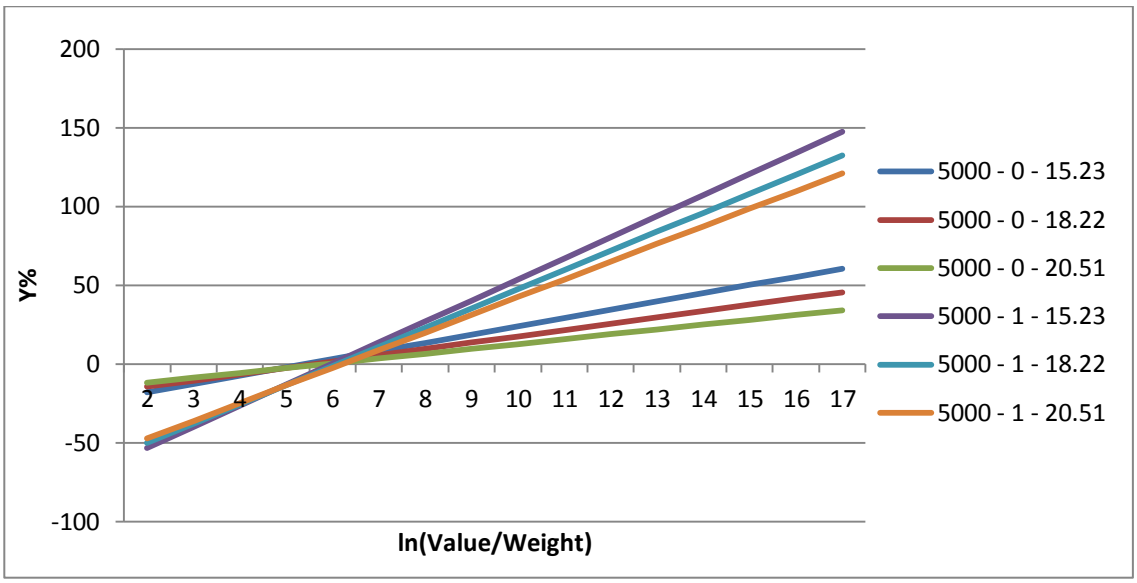
1000 Km



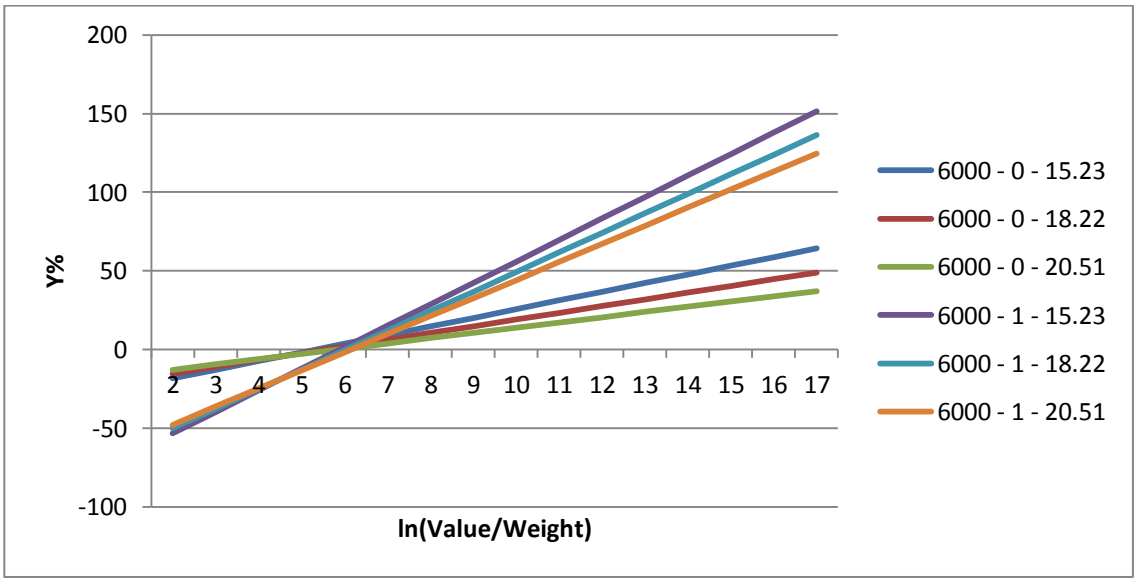
2500 Km



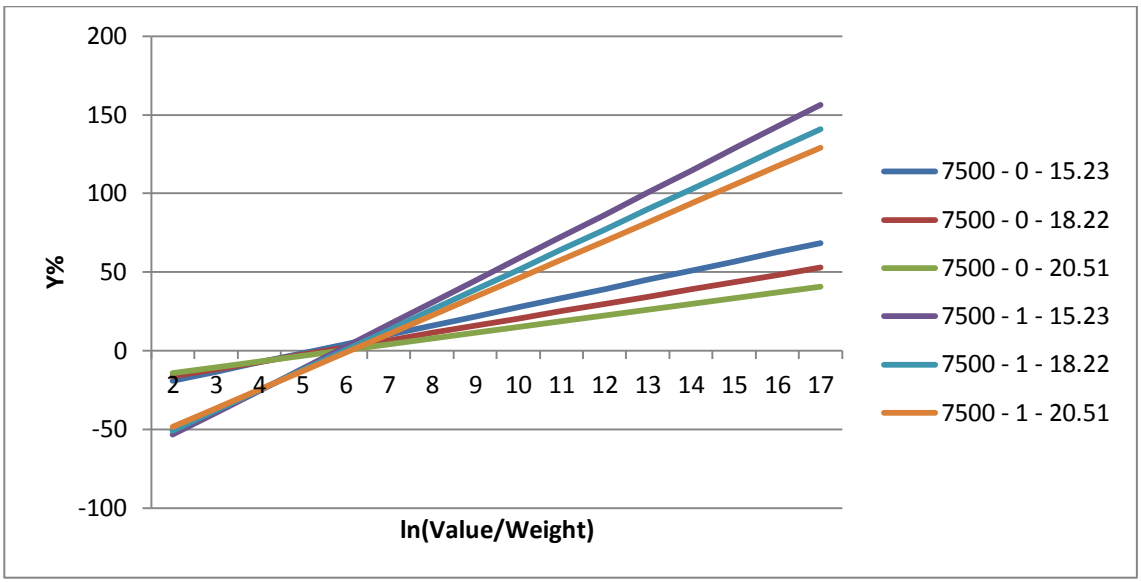
5000 Km



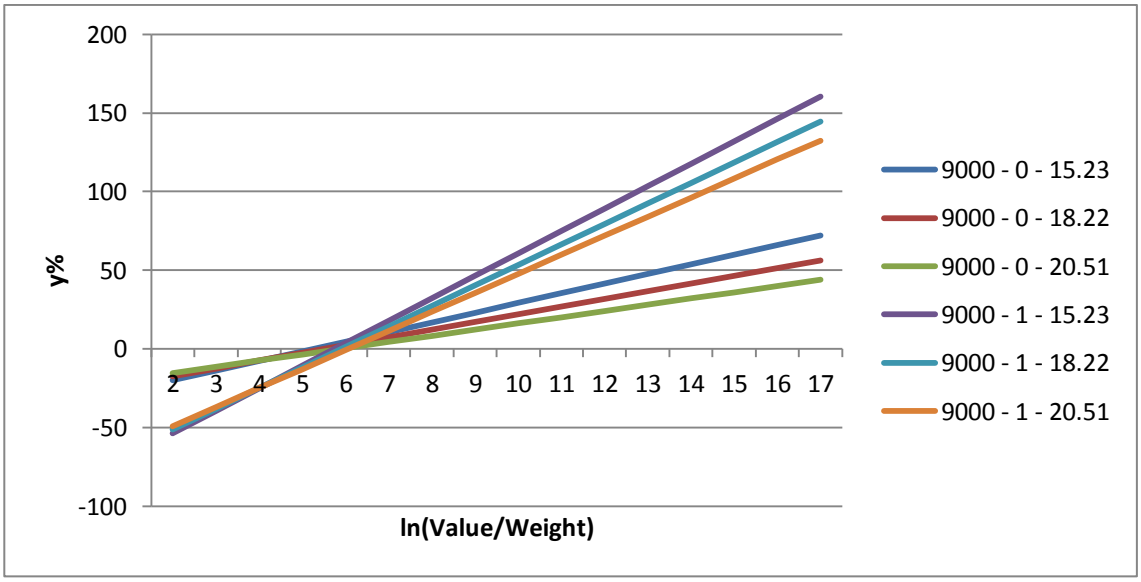
6000 Km



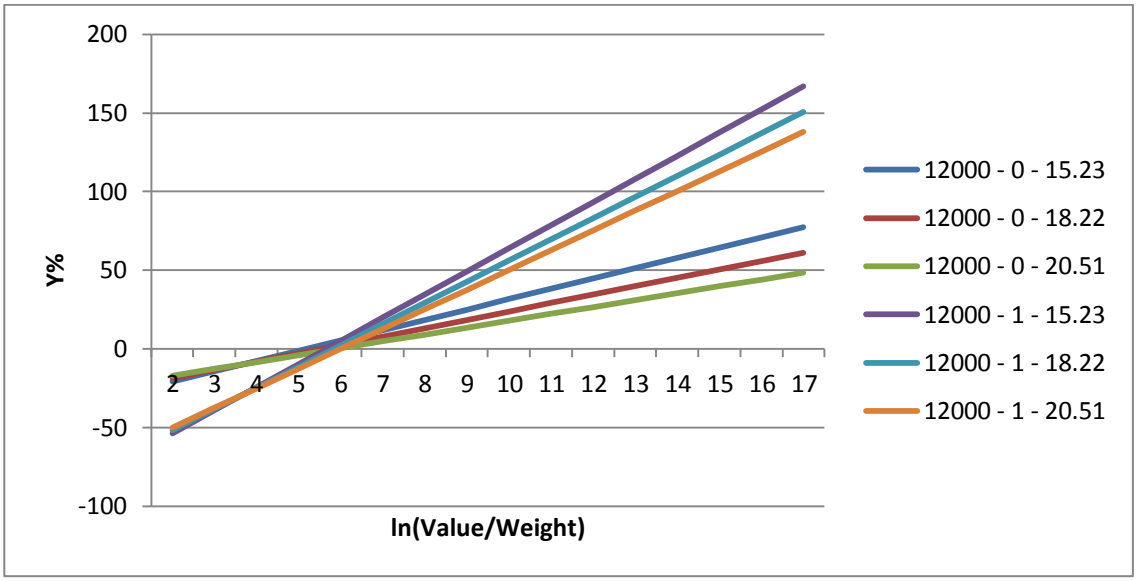
7500 Km



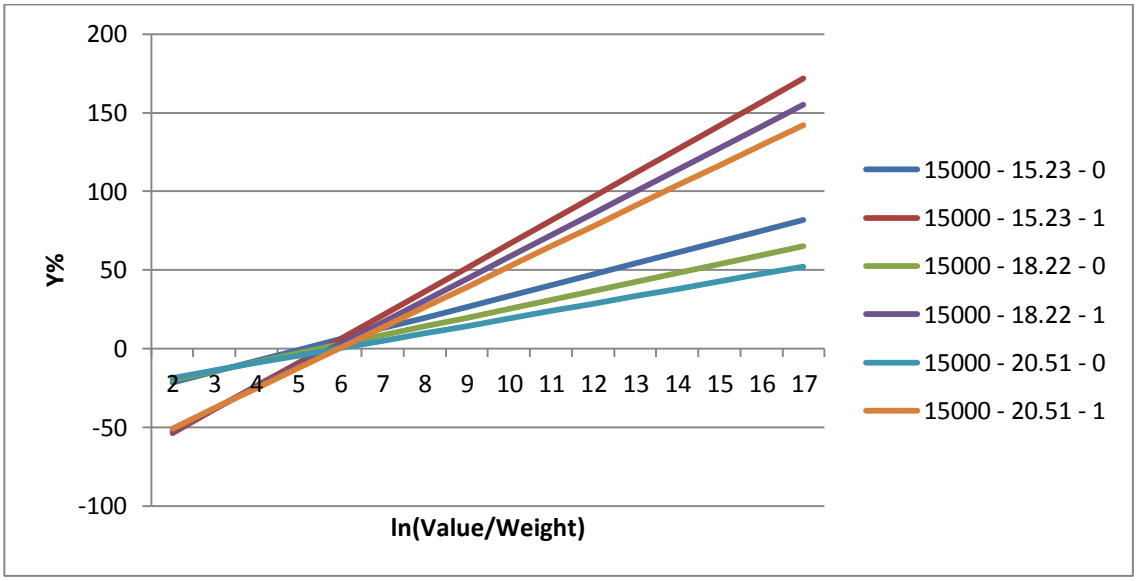
9000 Km



12000 Km



15000 Km



Variables Interdependences: Slacks Analysis

In order to individuate the most relevant influencing factor, we compare the different changes of "Y%" comparing two variables per time. Then we will "rank" the slacks (caused by the different variables variation) and consequently the variables impact.

Distance VS Value/Weight

	LowFare					MedianFare					HighFare				
	Low VW	Med VW	ΔY	High VW	ΔY	Low VW	Med VW	ΔY	High VW	ΔY	Low VW	Med. VW	ΔY	High VW	ΔY
Non-timely															
Low Dist	2.45	6.66	4.22	10.80	4.13	1.80	4.76	2.97	7.67	2.91	1.30	3.31	2.01	5.28	1.97
Med Dist	5.42	11.18	5.75	16.82	5.64	3.18	7.68	4.50	12.10	4.41	1.47	5.01	3.54	8.49	3.48
ΔY	2.98	4.51		6.02		1.38	2.92		4.43		0.17	1.70		3.21	
High Dist	6.38	12.63	6.25	18.76	6.13	3.63	8.63	5.00	13.53	4.90	1.52	5.56	4.04	9.52	3.96
ΔY	0.96	1.46		1.95		0.45	0.94		1.43		0.05	0.55		1.04	
Timely															
Low Dist	0.00	3.36	3.36	15.66	12.29	0.00	0.00	0.00	10.80	10.80	0.00	0.00	0.00	7.08	7.08
Med Dist	0.00	10.79	10.79	24.59	13.80	0.00	5.56	5.56	18.14	12.57	0.00	1.56	1.56	13.20	11.64
ΔY	0.00	7.43		8.93		0.00	5.56		7.34		0.00	1.56		6.12	
High Dist	0.00	13.19	13.19	27.48	14.29	0.00	7.45	7.45	20.51	13.06	0.00	3.05	3.05	15.18	12.12
ΔY	0.00	2.40		2.89		0.00	1.89		2.37		0.00	1.49		1.98	

Y is more sensitive to Value/Weight than to distance in average. For timely scenario Value/Weight becomes really influencing, this means there is strong "positive interdependency" between timeliness and Value/Weight: timely trade makes Y more sensitive to Value/Weight (it's clear even only looking at equation coefficient: 8,16). This can be interpreted in two ways. Firstly, if a time-sensitive

product requires fast delivery and the ideal partner is set far away, there are two main options: importing a similar product from closer country, or importing that product by air. The first option implies less "quality" since customers have to be content with products with only similar features, the second one implies extra-cost. As for any cost-quality trade off, decision depends on expected returns. Customers will engage in the expensive solution if it's very likely to generate high prices and returns (it usually implies very high Value/Weight or FOB prices). Secondly, Value/Weight and need for timeliness are related. Average Value/Weight for products with time constraint is much higher, so that the slacks analyzed here are in part misleading since if timely products have much higher Y and these products have in average higher Value/Weight, the higher Value/Weight implies higher predicted Y (but this because of need for timeliness, not because of higher Value/Weight).

Distance VS Fare3

	Low VW					Median VW					High VW				
	Low Fare	Med Fare	ΔY	High Fare	ΔY	Low Fare	Med Fare	ΔY	High Fare	ΔY	Low Fare	Med. Fare	ΔY	High Fare	ΔY
Non-timely															
Low Dist	2.45	1.80	-0.65	1.30	-0.50	6.66	4.76	-1.90	3.31	-1.45	10.80	7.67	-3.12	5.28	-2.39
Med Dist	5.42	3.18	-2.24	1.47	-1.72	11.18	7.68	-3.49	5.01	-2.67	16.82	12.10	-4.72	8.49	-3.61
ΔY	2.98	1.38		0.17		4.51	2.92		1.70		6.02	4.43		3.21	
High Dist	6.38	3.63	-2.75	1.52	-2.11	12.63	8.63	-4.01	5.56	-3.07	18.76	13.53	-5.23	9.52	-4.01
ΔY	0.96	0.45		0.05		1.46	0.94		0.55		1.95	1.43		1.04	
Timely															
Low Dist	0.00	0.00	0.00	0.00	0.00	3.36	0.00	-3.36	0.00	0.00	15.66	10.80	-4.86	7.08	-3.72
Med Dist	0.00	0.00	0.00	0.00	0.00	10.79	5.56	-5.22	1.56	-4.00	24.59	18.14	-6.45	13.20	-4.94
ΔY	0.00	0.00		0.00		7.43	5.56		1.56		8.93	7.34		6.12	
High Dist	0.00	0.00	0.00	0.00	0.00	13.19	7.45	-5.74	3.05	-4.40	27.48	20.51	-6.97	15.18	-5.33
ΔY	0.00	0.00		0.00		2.40	1.89		1.49		2.89	2.37		1.98	

Distance VS Timeliness

	Low VW			Med VW			High VW		
	0	1	ΔY	0	1	ΔY	0	1	ΔY
Low Fare									
LD	2.45	0.00	-2.45	6.66	3.36	-3.30	10.80	15.66	4.86
MD	5.42	0.00	-5.42	11.18	10.79	-0.39	16.82	24.59	7.77
ΔY	2.98	0.00		4.51	7.43		6.02	8.93	
HD	6.38	0.00	-6.38	12.63	13.19	0.56	18.76	27.48	8.72
ΔY	0.96	0.00		1.46	2.40		1.95	2.89	
Med Fare									
LD	1.80	0.00	-1.80	4.76	0.00	-4.76	7.67	10.80	3.13
MD	3.18	0.00	-3.18	7.68	5.56	-2.12	12.10	18.14	6.04
ΔY	1.38	0.00		2.92	5.56		4.43	7.34	
HD	3.63	0.00	-3.63	8.63	7.45	-1.18	13.53	20.51	6.98
ΔY	0.45	0.00		0.94	1.89		1.43	2.37	
High Fare									
LD	1.30	0.00	-1.30	3.31	0.00	-3.31	5.28	7.08	1.80
MD	1.47	0.00	-1.47	5.01	1.56	-3.45	8.49	13.20	4.71
ΔY	0.17	0.00		1.70	1.56		3.21	6.12	
HD	1.52	0.00	-1.52	5.56	3.05	-2.51	9.52	15.18	5.65
ΔY	0.05	0.00		0.55	1.49		1.04	1.98	

We can note again the strong interdependency between VW and Timeliness. For the Low VW scenario, time constraint implies negative slacks. Distance has stronger impact than timeliness for low VW while distance has much weaker impact (distance slacks are similar in all of the scenarios) for high VW. We can note the interdependencies between distance and VW, such as between distance and fare.

Value/Weight VS Fare3

	Low Dist					Med Dist					High Dist				
	Low Fare	Med Fare	ΔY	High Fare	ΔY	Low Fare	Med Fare	ΔY	High Fare	ΔY	Low Fare	Med Fare	ΔY	High Fare	ΔY
Non-Timely															
Low VW	2.45	1.80	-0.65	1.30	-0.50	5.42	3.18	-2.24	1.47	-1.72	6.38	3.63	-2.75	1.52	-2.11
Med VW	6.66	4.76	-1.90	3.31	-1.45	11.18	7.68	-3.49	5.01	-2.67	12.63	8.63	-4.01	5.56	-3.07
ΔY	4.22	2.97		2.01		5.75	4.50		3.54		6.25	5.00		4.04	
High VW	10.80	7.67	-3.12	5.28	-2.39	16.82	12.10	-4.72	8.49	-3.61	18.76	13.53	-5.23	9.52	-4.01
ΔY	4.13	2.91		1.97		5.64	4.41		3.48		6.13	4.90		3.96	
Timely															
Low VW	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Med VW	3.36	0.00	-3.36	0.00	0.00	10.79	5.56	-5.22	1.56	-4.00	13.19	7.45	-5.74	3.05	-4.40
ΔY	3.36	0.00		0.00		10.79	5.56		1.56		13.19	7.45		3.05	
High VW	15.66	10.80	-4.86	7.08	-3.72	24.59	18.14	-6.45	13.20	-4.94	27.48	20.51	-6.97	15.18	-5.33
DY	12.29	10.80		7.08		13.80	12.57		11.64		14.29	13.06		12.12	

In non-timely scenario VW and Fare have low similar impact over Y (slacks are not that different). We can note interdependencies among them and the other variables.

Timeliness VS Value/Weight

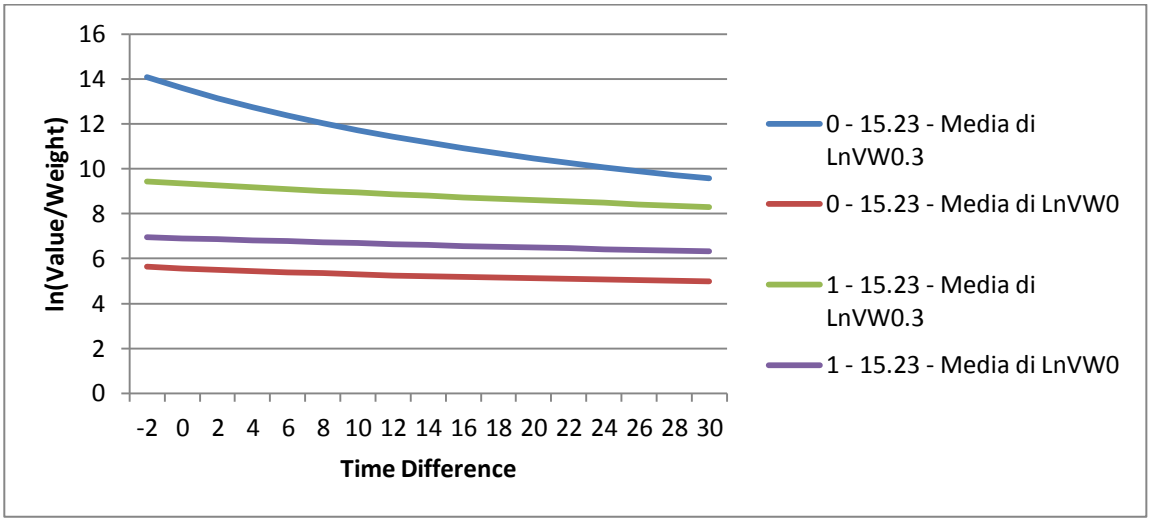
	Low Fare			Med Fare			High Fare		
	0	1	ΔY	0	1	ΔY	0	1	ΔY
Low Distance									
LVW	2.45	0.00	-2.45	1.80	0.00	-1.80	1.30	0.00	-1.30
MVW	6.66	3.36	-3.30	4.76	0.00	-4.76	3.31	0.00	-3.31
ΔY	4.22	3.36		2.97	0.00		2.01	0.00	
HVW	10.80	15.66	4.86	7.67	10.80	3.13	5.28	7.08	1.80
ΔY	4.13	12.29		2.91	10.80		1.97	7.08	
Med Distance									
LVW	5.42	0.00	-5.42	3.18	0.00	-3.18	1.47	0.00	-1.47
MVW	11.18	10.79	-0.39	7.68	5.56	-2.12	5.01	1.56	-3.45
ΔY	5.75	10.79		4.50	5.56		3.54	1.56	
HVW	16.82	24.59	7.77	12.10	18.14	6.04	8.49	13.20	4.71
ΔY	5.64	13.80		4.41	12.57		3.48	11.64	
High Distance									
LVW	6.38	0.00	-6.38	3.63	0.00	-3.63	1.52	0.00	-1.52
MVW	12.63	13.19	0.56	8.63	7.45	-1.18	5.56	3.05	-2.51
ΔY	6.25	13.19		5.00	7.45		4.04	3.05	
HVW	18.76	27.48	8.72	13.53	20.51	6.98	9.52	15.18	5.65
ΔY	6.13	14.29		4.90	13.06		3.96	12.12	

Timeliness VS Fare3

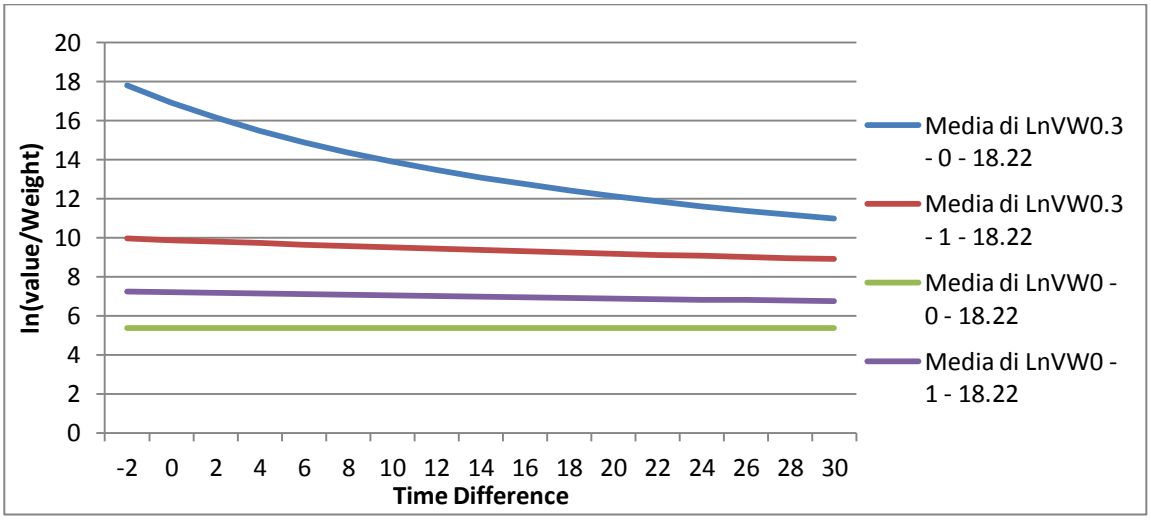
	Low VW			Med VW			High VW		
	0	1	ΔY	0	1	ΔY	0	1	ΔY
Low Distance									
LFare	2.45	0.00	-2.45	6.66	3.36	-3.30	10.80	15.66	4.86
MFare	1.80	0.00	-1.80	4.76	0.00	-4.76	7.67	10.80	3.13
ΔY	-0.65	0.00		-1.90	-3.36		-3.12	-4.86	
HFare	1.30	0.00	-1.30	3.31	0.00	-3.31	5.28	7.08	1.80
ΔY	-0.50	0.00		-1.45	0.00		-2.39	-3.72	
Med Distance									
LFare	5.42	0.00	-5.42	11.18	10.79	-0.39	16.82	24.59	7.77
MFare	3.18	0.00	-3.18	7.68	5.56	-2.12	12.10	18.14	6.04
ΔY	-2.24	0.00		-3.49	-5.22		-4.72	-6.45	
HFare	1.47	0.00	-1.47	5.01	1.56	-3.45	8.49	13.20	4.71
ΔY	-1.72	0.00		-2.67	-4.00		-3.61	-4.94	
High Distance									
LFare	6.38	0.00	-6.38	12.63	13.19	0.56	18.76	27.48	8.72
MFare	3.63	0.00	-3.63	8.63	7.45	-1.18	13.53	20.51	6.98
ΔY	-2.75	0.00		-4.01	-5.74		-5.23	-6.97	
HFare	1.52	0.00	-1.52	5.56	3.05	-2.51	9.52	15.18	5.65
ΔY	-2.11	0.00		-3.07	-4.40		-4.01	-5.33	

Variables Interdependences: The combined effect of Time in Framework

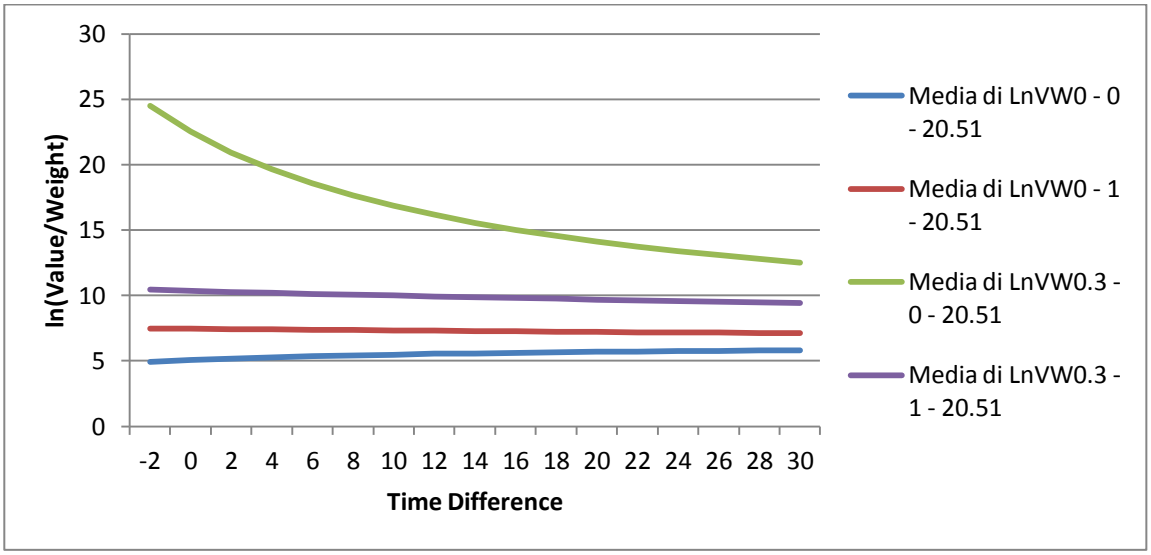
Low Fare



Median Fare



High Fare


30
25
20
15
10
5
0
ln(Value/Weight)
-2 0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30
Time Difference
Media di LnVW0 - 0 - 20.51
Media di LnVW0 - 1 - 20.51
Media di LnVW0.3 - 0 - 20.51
Media di LnVW0.3 - 1 - 20.51