The opportunities of the port of Constanza

Joost Piersma
400162

Abstract

To evaluate the opportunities of the port of Constanza we will investigate the effect of economical factors and improvements of infrastructure to the throughput and the amount of cargo transiting a country. We will focus on the South-East European region to link our findings to Romania and its competitors in the geographical area. We obtain significant differences between improvements on infrastructure and economical factors in low developed countries compared to high developed countries. Then we create a multinomial logit model to explore the incentives for shippers to select a specific port and following route by road, rail or barge transport to the hinterland regions. The most important findings derived from this model contain the negative and significant effect of transport time and the insignificant effect of transport costs. To investigate the efficiency of a port we include operational efficiency measures such as capital and labor productivity, and asset utilization rates to the multinomial logit model.

Contents

1 Problem description

This paper will analyze the opportunities of the port of Constanza to run the transport in South-East Europe in the future. The increasing container flows all over the world bring opportunities to well developed countries and perhaps even more for the less developed countries. South-East Europe could be considered as the relatively less developed countries of Europe. Opportunities could be for instance growth of the economy and economies of scale and scope. Transport can be divided in transshipment and gateway cargo. Transshipment corresponds to fully maritime transport and gateway cargo defines the incoming maritime transport, followed by road or railway transport. We will focus our regressions and model to a great extent on gateway cargo. First, the current position and situation of the port will be explored compared to the competitors in the South-East European region. This will be done according to existing literature research and data of hinterland connections, transport infrastructure, trade partners and throughput. The transport infrastructure will focus on the road, rail and barge transport. Next we will investigate the influence of some interesting variables related to the throughput of a set of countries and the amount of cargo transiting a set of countries. These influences will be explored by linear regressions. Then we will investigate the obtained and selected routes of shippers between the port and the following hinterland connections. We will look for incentives why shippers select a specific route by processing a multinomial logit model. This investigation will be based on the influences of interesting variables related to the market shares over the specific routes. Finally, we will focus on the efficiency of a port by adding operational efficiency measures such as capital and labor productivity, and asset utilization rates to the model. All these regressions and the multinomial logit model will be elaborated in sections 5 and 6.
The main research question of this paper is: *How could the port of Constanza improve and stimulate the economy of Romania?*
To elaborate this main question, we will extend the question with two subquestions: *Which factors stimulate the economy?* and *Which factors influence the port choice of shippers?*

2 Relevance

Further research on this topic will increase the information and opportunities of the economical value of the port of Constanza and her competitors. Between 1985 and 2000 the public research to this topic contained 37 papers, in contrast to over 200 papers between 2010 and 2014 (Neagoe, 2015). These facts show the increasing interest in this topic. Assuming that improvements of infrastructure in less developed counties do have a greater impact to the economy than developed countries, we think it is interesting to analyze what kind of infrastructure creates the most effect. Overall, this research will enable practical advice and improvements, and will contribute to a lesser extent to theoretical foundations.

3 Literature

To analyze the qualities and opportunities of the port of Constanza and her competitors, we will focus on the paper of Neagoe (2015). Most of the background information and used methods in this paper will be gathered from Neagoe (2015). To evaluate the current position of the port of Constanza, it is important to understand the motivation of a shipper to select a specific port. This choice could be substantiated by the following literature research. A multinomial logit model is created to understand the port choice of a shipper (Veldman and Buckmann, 2003). The South European ports do have an advantage over North European ports due to shorter travel time and distance for most of the container cargo from Asia, Africa and South America, but half of the cargo is handled in the ports of Le Havre-Hamburg (Notteboom, 2013). South European ports have started to develop the infrastructure and facilities, and have the advantage of lower costs to improve even more (Ferrari et al., 2006). A low level of privatization of the railway sector (Notteboom, 2010) and the lack of integration between economies (Medda and Carbonaro, 2007) in the South East European ports create some disadvantages. Especially the slow process of outsourcing the railway sector to the private sector hinders the speed and efficiency for rail transport (Cazzinga et al., 2002). Investments done by the private sector could improve this situation, in contrast cuts followed by governmental policies introduce the opposite effect. According to survey techniques the most important aspect of a port is the frequency of service (Slack, 1985), (Bird and Bland, 1988). But most of the time shippers are conservative in their choice for a port. Once they experience good treatment of service, they often remain to use this port (Slack, 1985). The most important elements for the shipping choices are often quality of service, infrastructure and connection (de Langen, 2007). Another approach is to analyze the situation introducing a multinomial logit model with variables as costs, and oceanic and inland distance included (Malchow, 2004). they conclude to a near-linear relation on the cost and transit time.

Next to the position, the impact of the port to the economy of a country is interesting to take in consideration. The connection between GDP and port throughput has been acknowledged as a dependent relation (Tsamourgelis et al., 2013). This argument could be substantiated by the reasoning that port throughput is a fraction of imports and exports, which are components of GDP. A higher level of cargo flows could be achieved by increasing transshipment cargo or gateway cargo. Supplying goods by inland waterway transport is a lucrative alternative (Rodrique et al., 2013). This result creates opportunities to focus on transshipment cargo. The ports of Piraeus and Ambarli locate themselves as main transshipment ports in the Mediterranean (container insight, 1 25). Ports are an important element of the transport supply chain, with the hinterland infrastructure of high importance (Tongzon, 2009). The economic performance of a country is correlated to the road cargo (McKinnon, 2007). When increasing the traffic handled by an expanded highway system would likely to improve the transport volume (Keeler and Ying, 1988), motivated by a reduction of the transport costs.

These results from previous research lead us to further investigations in most of the topics in this paper. We will analyze the relation and effect between variables to advice which elements would boost the economy in the South-East European region.

4 Motivation

Between 2008 and 2012 the major supplying ports of Constanza from the Mediterranean Sea, Piraeus and Ambarli, have increased their throughput volumes from China by 11 and 6 times respectively. In contrast to the container flows of Constanza, which declined with half of the production over that period. Piraeus and Ambarli increased their container flows by 7 times and 0.5 times, respectively, since 2008 (Neagoe, 2015). The container trade failure of the port of Constanza is probably one of the main reasons stagnation in the container flows occurs. On the other hand, these facts create opportunities by improving these amounts.

Next, the port of Constanza has access to the Danube, which provides possibilities to supply the transshipment cargo for Western Europe. The Danube could offer a cheap alternative for inland transport and create flexibility for shippers. According to an interview with Mr. Burgess (strategic research manager at Panteia) in Neagoe (2015) the port of Constanza does not make use of the full potential of the Danube. An explanation for that could be the unpredictable water level in the river during the year and the barge transport often requires quality in the last mile.

Trieste and Koper are the main competitors of the port of Constanza in the Adriatic range. These ports are mainly focused on gateway cargo. Many hinterland connections are provided by these ports. So, Constanza should focus on the hinterland connections which are poorly connected to these ports, or focus fully on the transshipment volumes and stay out of the com-

petition with these gateway cargo's.
The throughput volumes of the transshipment cargo of the port of Piraeus decreased from 0.9 million TEU in 2003 to 0.5 million TEU in 2007. The port of Constanza provided a transshipment cargo of more than 1 million TEU in 2008 from an amount of 0.15 million TEU in 2004 (Notteboom, 2013). These facts could motivate to focus even more on the transshipment cargo to create economies of scale and scope.
In the Northern European range ports there is no possibility to split the gateway and transshipment hubs, since they fulfill both purposes. Splitting these types of cargo could be an option for the Southern European range ports (Rodrique and Notteboom, 2010). To focus fully on the transshipment cargo, they could generate a higher scale with relatively lower prices per product. On the other hand, we should not forget that it will be a very hard task to overtake the, at the moment, better position of the port of Piraeus. Piraeus has a central position with respect to the Black Sea and Adriatic Sea, resulting in an advantage for travel distance for most of the incoming cargo. Furthermore they have a high level of terminal performance at their disposal. This quality can be substantiated by deep drafts, huge lengths of port berths and many quay cranes.
When making use of the handling of goods from the containers to a standard trailer for road network, these users experience relatively low labor costs at the port of Constanza. At the downside, they deal with a high premium over the standard kilometer cargo rate and some road network transport companies charge for the empty return trip.
All in all, these points lead, from our point of view, to a lot of opportunities which need to be investigated in more depth. With this research we could acquire more information about the strengths, weaknesses, opportunities and threats of the port of Constanza and her competitors.

5 Data

In this section we will firstly clarify the origin of the data used in the regressions and define the variables included in these regressions. Then we will analyze the transport time and transport cost to induct some practical implications. Next we will evaluate the selection procedure of the routes of the multinomial logit model. Finally we will clarify the characteristics related to the routes.

5.1 Data for regressions with dependent variable throughput

In this subsection we will introduce the explanatory variables included which could influence the dependent variable *throughput* of a country. The data for the regressions involving these variables coming up in section 6 is gathered from the research of Neagoe (2015), originally found in the EUROSTAT database. The data contains information from 24 European countries, with yearly observations from 2000 till 2013. Unfortunately not all the information over these 14 years is available for every country. The relative short time span could provide inconsistency in the estimation of the coefficients of the explanatory variables. The data contains information of the variables *throughput*, GDP, *trade*, *population* and information of binary variables. All the variables correspond to a certain country. The dependent variable *throughput* is the measured throughput of TEUs. The following variables are the explanatory variables. GDP is the market price measured in millions of euros. $Trade$ is modeled by a summation of *imports* and *exports*, which are goods measured in millions of euros. The first differences of *throughput* and GDP are obtained by the difference containing two following years. *Population* defines the population of a country in real values. The binary variables contain information to indicate if the explanatory variables are related to a certain country. The countries we will investigate with these binary variables are Romania, Bulgaria, Croatia, Greece and Slovenia.

5.2 Data for regressions with dependent variable transit

In this subsection we will introduce the explanatory variables included which could influence the dependent variable *transit* of a country. *Transit* is defined as the amount of cargo transiting a country. The following variables are the explanatory variables. The data for the regressions involving these variables coming up in section 6 is gathered from the research of Neagoe (2015) as well. Also originally found in the EUROSTAT database. The data contains information from 25 European countries, with yearly observations from 2003 till 2012. Unfortunately not all the information over these 10 years is available for every country. The even shorter time span than the data of the first six regressions could provide inconsistency in the estimation of the coefficients of the explanatory variables. The data contains information of the variables *transit*, *density*, *trade*, *GDP*, *population* and information of binary variables regarded as clusters. *Density* contains the highway density in kilometers per 100 square kilometers. *GDP* is the market price measured in millions of euros. *Trade* is modeled by a summation of *imports* and *exports*, which are goods measured in millions of euros. *Population* defines the population of a country in real values. The clusters contain sets of the 25 given countries of Europe ordered based on similar statistics of transit, density, trade, GDP and population. We use the same clusters as Neagoe (2015), except for Slovakia. Slovakia will be transfered to cluster 2 because of the most corresponding statistics to this cluster. This process will be clarified in more depth in section 7.

5.3 Data for multinomial logit model

This subsection will contain the evaluation of the data used in the multinomial logit model. The multinomial logit model will be used to investigate the choice of a shipper for a specific route. This choice will be expressed in observed cargo flows of transport over the selected routes. The choice of this selected route will be influenced by its characteristics for instance based on the qualities of a port, the cost of the transport and the time of the transport. Later on in this subsection we will go further in detail of the data of the multinomial logit model, especially the selection of the routes included in the model. First we will take a look at the choice based on the transport time. Doing this by analyzing the average speed corresponding to a specific route of the transport by road, rail and barge. The data we use for this analysis is gathered from the ETISplus database and contains the 2203 possible routes for road, rail and barge transport. Later on we will describe the selection procedure to find these 2203 routes. Possibly this could clarify the choice of a shipper for a specific route as well.
In the following figure we obtain the average speed (y-axes) with the corresponding distances (x-axes) for the road transport (1) from our dataset. We observe a range of distance of 0 to 2000 kilometers, and a range of average speed of 0 to 80 kilometers per hour. The 1078 observations give a sight of the combinations of distance and average speed.

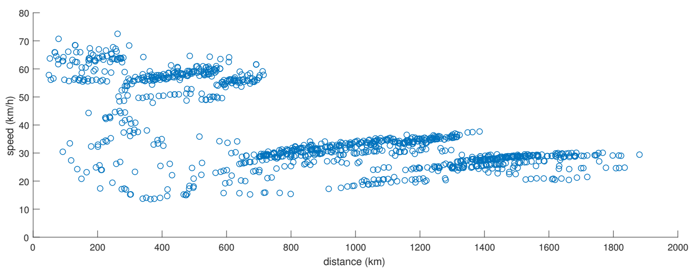


Figure 1: Average speed of road transport

The first point which strikes us is the drop of the average speed for distances over 700 kilometers.

According to Neagoe (2015) this is probably caused by international regulations which oblige truck drivers to take a sleeping brake after driving for 10.5 hours. Planning the sort of transport for a specific route with corresponding distance could be influenced by taking this reasoning in consideration. Routes with a range under 700 kilometers could be relatively more efficient than routes with a range over 700 kilometers.
In the next figure we obtain the average speed (y-axes) with the corresponding distances (x-axes) for the barge transport (2) from our dataset. We observe a range of distance of 0 to 1600 kilometers, and a range of average speed of 6.5 to 7.1 kilometers per hour. There are 49 observations available. The small number of observations could provide a misspecification in our evaluation. Furthermore, take into account that all the routes of the observations start at the port of Constanza.

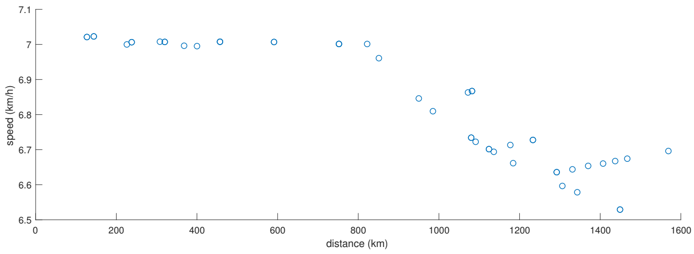


Figure 2: Average speed of barge transport

We observe a steady average speed for barge transport up to 800 kilometers from around 7 kilometers per hour. After the boundary point of this range we observe a near linear decrease of the average speed. According to Neagoe (2015) this decline is observed because of the waiting time to pass the locks at the Iron Gates power station and other locks of the Danube.
In the last figure of this subsection, we observe the average speed (y-axes) with the corresponding distances (x-axes) for the rail transport (1) from our dataset. We observe a range of distance of 0 to 2000 kilometers, and a range of average speed of 0 to 60 kilometers per hour. There are 1076 observations available.

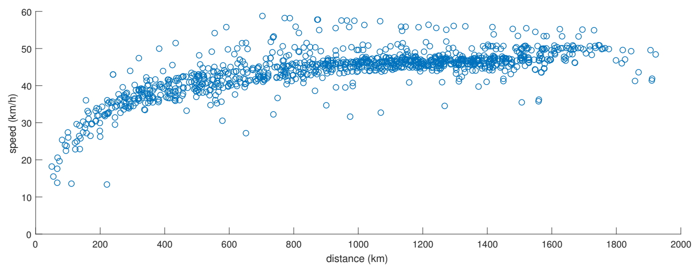


Figure 3: Average speed of rail transport

We observe an increasing trend for the average speed for rail transport for a distance within the range of the first 800 kilometers. After passing the distance of 800 kilometers the average speed stabilizes around 60 kilometers per hour. Combining these results of the observations with a small variance of the observations per distance, we state there are not many restrictions and problems for rail transport.
Overall, when we take the possibilities for the sort of transport in consideration to make the process time efficient per distance we come to the following results. When we need to transport goods and products within a range of 700 kilometers it is mostly time efficient to select road transport. For distances over this range it would be more time efficient to select rail transport. Barge transport would certainly not be a time efficient alternative, but could be an alternative with an advantage of low costs. The costs of the alternatives will be discussed in this subsection

as well. In the table below we observe significant results at a 1% significance level for the coefficients of the explanatory variable *distance* to the dependent variable *average speed*. These results statistically confirm the negative effect of *distance* to *average speed* for road and barge transport and the positive effect of *distance* to *average speed* for rail transport.

Table 1: Regression results with dependent variable average speed

Coefficient	Road	Barge	Rail
Constant	54.8492	7.1330	34.7389
T-statistic	76.88**	323.69**	110.77**
Distance	-0.0199	-0.0003	0.0096
T-statistic	-28.08**	-15.16**	32.35**
R-squared	0.4229	0.8303	0.4936
Adj. r-squared	0.4224	0.8266	0.4931
Observations	1078	49	1076

*, ** denote the coefficients are significant at a 5%, 1% significance level respectively. Adj. r-squared defines the adjusted r-squared.

We suggest that rail transport would be the cheapest alternative for every route and every size of containers. To investigate the alternatives of transport we will focus on rail and barge transport. These sorts of transport are more comparable because of the number of containers they can transport per trip in contrast to road transport. The data we use is modeled by the research of Neagoe (2015). We evaluate the average costs in euros per kilometer per container size. The container sizes included are 20ft., 40 ft. and 45ft.. The costs are generalized to 20ft. containers, so the average costs of 40ft. containers are divided by 2 and the average costs of 45ft. containers are divided by 2.25. In the table below we observe that rail is not in every situation the cheapest alternative, thus the alternatives should be evaluated for every specific route when we focus on cost efficiency.

Table 2: Costs

Sort of transport	20ft.	40ft.	45 ft.
Barge	0.48	0.41	0.40
Rail	0.52	0.36	0.33

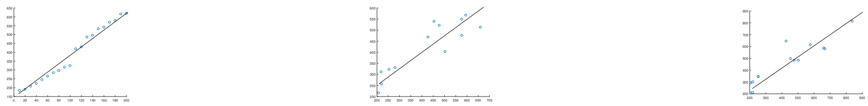
20ft., 40ft. and 45ft. are the different sizes of the containers. The values below the sizes are the costs in euros per kilometer per container size generalized to 20ft. containers to ease the evaluation and corresponding sort of transport.

Now we will focus on the data used in the multinomial logit model. To process the multinomial logit model we include explanatory variables and binary variables which are related to a specific route (route m) and the base route (route b). These variables could influence the choice of the selected route. The routes we will analyze contain transport between one of the eight ports and the hinterland regions of Bulgaria, Romania, Hungary, Croatia, Slovenia, Greece, Serbia and Macedonia (route m), and transport between the port of Constanza and the hinterland regions of the just mentioned eight countries (route b). We will investigate the effects adjusted to route m with the corresponding ports which are Burgas, Varna, Thessaloniki, Piraeus, Rijeka, Trieste, Koper and Constanza as well. We selected these ports because they handle over 20.000 TEU per year (Neagoe, 2015). The characteristics of route m contain information corresponding to this route and the port in each specific case, and the characteristics of route b contain information corresponding to this route and the port of Constanza. We divide the countries into NUTS 3 regions, and will focus on the routes connecting the eight ports to these regions. This classification is based on the population per region. NUTS 1 regions have between 3 and 7 million people, NUTS 2 regions have between 0.8 and 3 million people and NUTS 3 regions have between .15 and .8 million people. By selecting these NUTS 3 regions we focus on the most detailed routes. Before we will clarify which variables are included in the multinomial logit model, we model the routes we want to investigate with the multinomial logit model.

We model these routes separately for every country and divide the routes over road, rail and barge transport to avoid small mistakes in the programming code and importing the data. This procedure results in more interaction with the data, which at the downside requires more time. All the selection procedures related to the routes are provided by MATLAB. We extract all the possible routes by road, rail and barge transport to the hinterland regions of the eight given countries from the ETISplus (2014) database. Then we eliminate all the routes which do not originate in one of the eight ports and we eliminate the routes from one of the eight ports to one of the other seven ports, because we aim to focus on the routes to the hinterland regions. Finally, we eliminate the routes which are reachable from one or more of the ports, but not reachable from the port of Constanza. We make these decisions, because we will evaluate route m from one of the eight given ports to the hinterland versus the base route to this same hinterland destination, so we need the distance of the base route. After this selection procedure we obtain 2203 possible routes, which contain 1078 routes for road transport, 49 routes for barge transport and 1076 routes for rail transport.

Next, we will extract all the routes with cargo flows of road, rail and barge transport to the hinterland regions of the eight given countries from the ETISplus (2014) database. The dependent variable of the multinomial logit model is based on the relative market share of the selected route compared to the relative market share of the base route. The selected route m could be a specific route, but could be the base route as well. For this reason we need to investigate the cargo flows of the routes. We also eliminate all the routes which do not originate in one of the eight ports and we eliminate the routes from one of the eight ports to one of the other seven ports. Then we collect the total cargo flow from a port to a specific hinterland region, because often there are multiple cargo flows observed. This selection procedure provides 920 routes with cargo flows, which contain 667 routes for road transport, 23 routes for barge transport and 230 routes for rail transport. Now we might consider two methods of modeling the data for the fractions of cargo flows per route. The first method collects the total cargo flows to every hinterland region. After collecting these cargo flows from the port to the hinterland region we determine the relative market share of this route by dividing this cargo flows by the total cargo flows to this specific hinterland region. Finally, we link these fractions of cargo flows of the specific routes with the already gathered 2203 possible routes. The other method first links the cargo flows per specific route to the already gathered 2203 routes. Then collects the total cargo flows to every hinterland region. Finally, we determine the relative market share of the cargo flows of the specific route by dividing the cargo flows of this route by the total cargo flow to this hinterland region. These methods provide different results, because not every route that is available in the cargo flows file is available in the possible routes file, and vice versa. Combining these two files provide 850 routes, which contain 657 routes for road transport, 172 routes for rail transport and 21 routes for barge transport. We evaluate both methods and decide to process the first method, because of over valuation of the relative market share of the cargo flows in the second method, which results in misspecification of the model. In accordance with Neagoe (2015) we will determine the base market share as the fraction of cargo flows from the port of Constanza to the specific hinterland region compared to the total cargo flows to this hinterland region. Furthermore, if there is no route obtained with cargo flows from the port of Constanza to the hinterland region we determine the base market share with a value of 0.0001. This procedure allows us to include this route to the multinomial logit model. This relatively low market share provides a small value of utility for the shipper to select the base route. To treat the data fairly this procedure introduces a base market share limit of 0.0001.

Now we will focus on the explanatory variables which could influence the motivation for the selected route. The *transport time* will be defined as the observed minutes per route, which is gathered from the ETISplus (2014) database when extracting the possible routes. The *transport costs* are gathered from the research of Neagoe (2015). This information contains *transport costs* for barge transport between 207 and 660 kilometers, for rail transport between 207 and 838 kilometers and for road transport between 10 and 200 kilometers. Barge and rail transport contain the *transport costs* for different sizes of containers. 20 ft. containers are defined as 1 TEU. To ease the transport costs, we generalize all the costs to 1 TEU by dividing the 40ft. and 45ft. by 2 and 2.25 respectively. as we can see in the scatter plots of the data (figure 4) below we notice a near linear relation between the transport costs and the transport distance.



(a) road (b) barge (c) rail

Figure 4: Scatter plots of costs versus distance

To analyze the transport costs we take an average of the generalized costs. We regress the transport distance to the *transport costs* to estimate the *transport costs* for transport routes which contain a distance out of the available range of distances included in the data of Neagoe (2015), but included in the possible routes. To estimate these *transport costs* we create a function including the estimated coefficient of the constant and the transport distance. This function generates a difference for the values in the available and possible routes of at most 20% for road transport, at most 30% for barge transport and at most 40% for rail transport. On average these amounts are more within reason with a difference of 6%, 12% and 11% for road transport, barge transport and rail transport respectively. The *maritime diversion distance* will be gathered from Notteboom and Dooms (2014). The *maritime diversion distance* contains the nautical miles distance from the port corresponding to the route to the main trade lane. The *maritime diversion distance* are obtained for Constanza, Thessaloniki, Piraeus, Trieste and Koper. In accordance with Neagoe (2015) we estimate the *maritime diversion distance* for the other ports, Burgas, Varna and Trieste, by adding the distance in nautical miles between these ports and the nearest port available in Notteboom and Dooms (2014) separately. This port-to-port nautical miles distance is gathered from sea distances (2018). The *crane congestion* will be defined as the throughput of the port in TEU in 2013 divided by the number of gantry cranes available at the corresponding terminal. This information is gathered from Neagoe (2015), except for the information of the gantry cranes available in the port of Burgas. According to port of Burgas (2017) there are 2 gantry cranes available in the port of Burgas and already available in 2013. For every port the terminal handling charge (THC), maximum port draft in meters (*port depth*) and available berthing space in meters (*port berth*) are gathered from Neagoe (2015) as well. According to Neagoe (2015) there are two terminals available for Piraeus, Varna and Constanza. To determine the *port depth* we generate the weighted averages of the nominal capacity handled in the corresponding terminals. Combining these weighted averages to the statistics of each terminal for *port depth* generates the average characteristic of each corresponding port. To determine the *port berth* we add up the lengths of the two terminals. After determine the possible routes for rail and barge transport we will generate the binary vector when each of these methods of transport is used, D_{barge} and D_{rail}. By checking if the country of the destination of the route is the same as the country of the port where the transport originates from we can set up the binary vector $\mathrm{D}_{same\,country\,shipment}$. To create the binary variables, $\mathrm{D}_{ownership(shipping\,line)}$ and $\mathrm{D}_{ownership(terminal\,op)}$, corresponding to the terminal ownership and management we used information gathered from Neagoe (2015).

5.4 Data for multinomial logit model - operational efficiency measures

This subsection will evaluate the data used in the expanded model, the multinomial logit model - operational efficiency measures. According to Tongzon and Ganesalingam (1994) we expand the model with operational efficiency measures. These measures contain capital and labor productivity, and asset utilization rates. To include capital and labor productivity to the model we suggest the ratio of TEUs handled in the port divided by the number of employees of the port (*employee rate*) and the ratio of the number of TEUs handled in the port divided by the number of available tugs of the port (*tug rate*). To include asset utilization rates to the model we suggest the ratio of the number of TEUs handled in the port divided by the length of the berth of the port (*berth rate*) and the ratio of the number of TEUs handled in the port divided by the size of the terminal area of the port (*terminal rate*). The TEUs handled in the corresponding port are gathered from the research of Neagoe (2015) and come from 2013. The data of the employees of the port is mainly gathered from the research of van Hooydonk (2013) and contains information varying from 2009 up to and including 2012. We estimate the number of employees of the corresponding port in 2013 by including the industrial yearly growth of the country of the corresponding port according to the information gathered from eurostat (2018). Number of employees in the ports of Trieste and Koper are gathered from the corresponding port authority page which are up to date. To estimate the number of employees in 2013 we discount these numbers by the industrial yearly growth of the country of the corresponding port gathered from OECD (2018) for the port of Trieste and tradingeconomics (2018) for the port of Koper. The data of the size of the terminal and the available tugs per port are gathered from the pages of the corresponding port authorities as well. Some of these observations are not available on the port authority's page and is gathered from global agents, such as harbours review (2018b). These values are up to date, so we investigated the annual reports of the corresponding ports for changes in these values between 2013 and 2018. We did not found significant and finalized plans to the corresponding ports, thus we assumed them to process these findings. The data for the length of berth corresponding to the port used for the *berth rate* is gathered from the research of Neagoe (2015). The data of the nested variables of the original multinomial logit model are gathered from the research of Neagoe (2015) as well.

6 Methodology

In this section we will clarify the regressions to investigate the effect of some interesting variables to the dependent variables *throughput* and *transit*. As mentioned in section 5, *throughput* is the measured throughput of TEUs and *transit* is defined as the amount of cargo transiting the country. Next we will introduce the multinomial logit model. According to this model, we will look for incentives of a shipper to select route m compared to the base route. Route m is the route between one of the eight ports, described in section 5, to the hinterland region and the base route is the route between the port of Constanza and the hinterland region.

6.1 Regressions with dependent variable throughput

The first linear regression is deduced by the research of Tsamourgelis et al. (2013). This research found a positive and significant relation between GDP and *throughput*. The regression expands from the relation between the explanatory variable GDP and dependent variable *throughput* (1) to a logarithmic transformation including extra explanatory variables *trade* and *population*, and some binary variables (6). The logarithmic transformation is possible, because we are dealing with positive values. Furthermore, the logarithmic scale could provide less variation in the error terms caused by more close to normal distributed error terms. The relation between the first differences of GDP and *throughput* will be evaluated as well. The binary variables are linked to Romania, Bulgaria, Croatia, Greece and Slovenia because these countries contain the main competitors of the port of Constanza, including Constanza itself, and are geographically scaled South-East Europe. According to Neagoe (2015) Italy has a relatively low share of the contribution of throughput and is kept out of the regression. Equation 6 will represent the effect of the explanatory variables to the dependent variable compared to the other 19 European countries.

$$\text{Throughput} = \alpha + \beta \times \text{GDP} + \varepsilon. \tag{1}$$

$$\Delta\text{Throughput} = \alpha + \beta \times \Delta\text{GDP} + \varepsilon. \tag{2}$$

$$\ln(\text{Throughput}) = \alpha + \beta \times \ln(\text{GDP}) + \varepsilon. \tag{3}$$

$$\ln(\text{Throughput}) = \alpha + \beta_1 \times \ln(\text{GDP}) + \beta_2 \times \ln(\text{Trade}) + \varepsilon. \tag{4}$$

$$\ln(\text{Throughput}) = \alpha + \beta_1 \times \ln(\text{GDP}) + \beta_2 \times \ln(\text{Trade}) + \beta_3 \times \ln(\text{Population}) + \varepsilon. \tag{5}$$

$$\begin{aligned}\ln(\text{Throughput}) = \alpha + \beta_1 \times \ln(\text{GDP}) + \beta_2 \times \ln(\text{Trade}) + \beta_3 \times \ln(\text{Population}) + \beta_4 \times \text{D}_{Rom} \\ + \beta_5 \times \text{D}_{Bul} + \beta_6 \times \text{D}_{Cro} + \beta_7 \times \text{D}_{Gre} + \beta_8 \times \text{D}_{Slo} + \varepsilon,\end{aligned} \tag{6}$$

where D_{Rom}, D_{Bul}, D_{Cro}, D_{Gre} and D_{Slo} denote Romania, Bulgaria, Croatia, Greece and Slovenia, respectively. If an observation refers to one of these countries, it will contain the value 1, otherwise 0.

6.2 Regressions with dependent variable transit

According to the work of Keeler and Ying (1988) the relation between transit cargo and highway density (7) is created. This linear regression will be evaluated by the explanatory variable *density* to the dependent variable *transit*. This regression expands to a logarithmic transformation including extra explanatory variables *trade*, *GDP*, *population* and cluster variables (13). Again this logarithmic transformation is possible because of the positive values and will reduce the variance in the error terms compared to the regular scale. Clusters will be introduced, because we suggest there is a difference in effect by improving the infrastructure of high or less developed countries. Transit cargo, highway density, trade, GDP and population were included as standards to determine the clusters. Cluster 2 contains countries with the smallest economies of Europe, cluster 3 contains countries with the largest economies of Europe, cluster 1 contains the rest of countries of Europe. We are looking for positive effects after improving the infrastructure of a cluster, thus the final model (13) will contain every cluster that has a positive significant effect on the dependent variable *transit*. We can not include all clusters, in that case one regressor can be written as a linear combination of the other regressors. The regressors will be exactly collinear.

$$\text{Transit} = \alpha + \beta \times \text{Density} + \varepsilon. \tag{7}$$

$$\ln(\text{Transit}) = \alpha + \beta \times \ln(\text{Density}) + \varepsilon. \tag{8}$$

$$\ln(\text{Transit}) = \alpha + \beta_1 \times \ln(\text{Density}) + \beta_2 \times \ln(\text{Trade}) + \varepsilon. \tag{9}$$

$$\ln(\text{Transit}) = \alpha + \beta_1 \times \ln(\text{Density}) + \beta_2 \times \ln(\text{Trade}) + \beta_3 \times \ln(\text{GDP}) + \varepsilon. \tag{10}$$

$$\ln(\text{Transit}) = \alpha + \beta_1 \times \ln(\text{Density}) + \beta_2 \times \ln(\text{Trade}) + \beta_3 \times \ln(\text{GDP}) + \beta_4 \times \ln(\text{Population}) + \varepsilon. \tag{11}$$

$$\begin{aligned} \ln(\text{Transit}) = \alpha + \beta_1 \times \ln(\text{Density}) + \beta_2 \times \ln(\text{Trade}) + \beta_3 \times \ln(\text{GDP}) + \beta_4 \times \ln(\text{Population}) \\ + \beta_5 \times \text{Cluster i} + \varepsilon, \end{aligned} \tag{12}$$

where $\mathrm{i} \in \{1, 2, 3\}$.

$$\begin{aligned} \ln(\text{Transit}) = \alpha + \beta_1 \times \ln(\text{Density}) + \beta_2 \times \ln(\text{Trade}) + \beta_3 \times \ln(\text{GDP}) + \beta_4 \times \ln(\text{Population}) \\ + \beta_5 \times \text{Cluster 1} \times \mathrm{I}_1 + \beta_6 \times \text{Cluster 2} \times \mathrm{I}_2 + \beta_7 \times \text{Cluster 3} \times \mathrm{I}_3 + \varepsilon, \end{aligned} \tag{13}$$

where I_1, I_2 and I_3 will contain the value 1, if the corresponding cluster 1, cluster 2 and cluster 3 respectively was evaluated in equation 12 with a positive and significant coefficient, otherwise 0.

6.3 Multinomial logit model

Veldman and Buckmann (2003) defined a multinomial logit model to look for the incentives for a specific route for shippers. Set route m as a specific route from one of the eight included ports to a hinterland region. The fundamentals of the multinomial logit model are based on the transport costs (C_m), transport time (T_m) and maritime deviation distance (M_m) of route m. These components create a utility function (14) for the shipper. This utility function creates value according to rationality, in contrast to the most of the micro economic utility functions. We will evaluate the utility of route m as the fraction of cargo flows to a specific hinterland region over this route m. Thus according to the utility function we investigate the influence of the variables to the observed cargo flows.

$$\mathrm{U}_m = \beta_1 \times \mathrm{C}_m + \beta_2 \times \mathrm{T}_m + \beta_3 \times \mathrm{M}_m. \tag{14}$$

The probability of selecting route m from the set of all possible routes r is defined by the probability of cargo flows (15).

$$\mathrm{P}_m(m = r | r = 1...M) = \frac{\exp(\mathrm{U}_m)}{\sum_{r=1}^{r=M} \exp(\mathrm{U}_r)}. \tag{15}$$

Set route b as the base route from the port of Constanza to a hinterland region. We could interpret this base route as an identification of the multinomial logit model. Identification is necessary to create a unique model. This route from the port to the hinterland region will be covered by road, rail or barge transport. Now the shipper can compare route m against the base route. This comparison can be seen as an odds ratio (16). To express this choice in measurable variables, we compare the relative market shares of cargo flows of route m to the relative market shares of cargo flows of route b (base market share). This trade-off will be the dependent variable in the multinomial logit model.

$$\frac{\mathrm{P}_m}{\mathrm{P}_b} = \frac{\exp(\mathrm{U}_m)}{\exp(\mathrm{U}_b)} = \exp(\mathrm{U}_m - \mathrm{U}_b). \tag{16}$$

When linking the utility function (14) to the probability function (16), we could create a model which evaluates the effects of the explanatory variables of the utility function to the choice of a shipper for selecting route m compared to the base route. A logarithmic transformation of this model provides more sense for interpreting the coefficients of these variables and is possible because of the positive values. This transformation provides less variation in the error terms which will be more close to normal distributed, resulting in the following model:

$$\ln\left(\frac{\mathrm{P}_m}{\mathrm{P}_b}\right) = \ln\left(\frac{\exp(\mathrm{U}_m)}{\exp(\mathrm{U}_b)}\right) = \beta_1 \times (\mathrm{C}_m - \mathrm{C}_b) + \beta_2 \times (\mathrm{T}_m - \mathrm{T}_b) + \beta_3 \times (\mathrm{M}_m - \mathrm{M}_b). \tag{17}$$

Equation (17) will introduce the multinomial logit model and could be expanded with some interesting variables and binary variables to increase the fit of the model. A better fit improves the reliability of the interpretation of the results. Including the constant α is a striking difference compared to the multinomial logit model of Neagoe (2015). After processing equation (17) we notice this procedure to be necessary. This evaluation will be elaborated in section 7. The variables included in the desired model evaluates the characteristics of route m compared to the characteristics of the base route. These variables are, as mentioned in section 5, the throughput divided by the number of gantry cranes available at the terminal (*Crane congestion*), the terminal handling charge (*THC*), the maximum port draft (*Port depth*) and the available port berths (*Port berth*). The binary variables involved evaluate the effect of the explanatory variables if rail transport is used (D_{rail}), if barge transport is used (D_{barge}), if the origin and destination of the route are in the same country ($\mathrm{D}_{same\,country\,shipment}$), if the terminal used is owned by a shipping line ($\mathrm{D}_{ownership(shipping\,line)}$) and if the terminal used is owned by a terminal operator ($\mathrm{D}_{ownership(terminal\,op)}$). The binary variables will contain a 1 if included in route m, otherwise 0.

$$\ln\left(\frac{\mathrm{P}_m}{\mathrm{P}_b}\right) = \alpha + \beta_1 \times \Delta\mathrm{C}_m + \beta_2 \times \Delta\mathrm{T}_m + \beta_3 \times \Delta\mathrm{M}_m + \varepsilon. \tag{18}$$

$$\ln\left(\frac{\mathrm{P}_m}{\mathrm{P}_b}\right) = \alpha + \beta_1 \times \Delta\mathrm{C}_m + \beta_2 \times \Delta\mathrm{T}_m + \beta_3 \times \Delta\mathrm{M}_m + \beta_4 \times \mathrm{D}_{rail} + \beta_5 \times \mathrm{D}_{barge} + \varepsilon. \tag{19}$$

$$\begin{aligned}\ln\left(\frac{\mathrm{P}_m}{\mathrm{P}_b}\right) = \alpha &+ \beta_1 \times \Delta\mathrm{C}_m + \beta_2 \times \Delta\mathrm{T}_m + \beta_3 \times \Delta\mathrm{M}_m + \beta_4 \times \mathrm{D}_{rail} + \beta_5 \times \mathrm{D}_{barge} \\ &+ \beta_6 \times \mathrm{D}_{same\,country\,shipment} + \varepsilon.\end{aligned} \tag{20}$$

$$\begin{aligned}\ln\left(\frac{\mathrm{P}_m}{\mathrm{P}_b}\right) = \alpha &+ \beta_1 \times \Delta\mathrm{C}_m + \beta_2 \times \Delta\mathrm{T}_m + \beta_3 \times \Delta\mathrm{M}_m + \beta_4 \times \mathrm{D}_{rail} + \beta_5 \times \mathrm{D}_{barge} \\ &+ \beta_6 \times \mathrm{D}_{same\,country\,shipment} + \beta_7 \times \mathrm{D}_{ownership(shipping\,line)} \\ &+ \beta_8 \times \mathrm{D}_{ownership(terminal\,op)} + \varepsilon.\end{aligned} \tag{21}$$

$$\begin{aligned}\ln\left(\frac{\mathrm{P}_m}{\mathrm{P}_b}\right) = \alpha &+ \beta_1 \times \Delta\mathrm{C}_m + \beta_2 \times \Delta\mathrm{T}_m + \beta_3 \times \Delta\mathrm{M}_m + \beta_4 \times \mathrm{D}_{rail} + \beta_5 \times \mathrm{D}_{barge} \\ &+ \beta_6 \times \mathrm{D}_{same\,country\,shipment} + \beta_7 \times \mathrm{D}_{ownership(shipping\,line)} \\ &+\beta_8 \times \mathrm{D}_{ownership(terminal\,op)} + \beta_9 \times \Delta\text{Crane congestion} + \beta_{10} \times \Delta\text{THC} \\ &+ \beta_{11} \times \Delta\text{Port berths} + \beta_{12} \times \Delta\text{Port depth} + \varepsilon,\end{aligned} \tag{22}$$

where Δ will indicate the difference of the characteristics of specific route m compared to the characteristics of base route b. For example, $\Delta\mathrm{C}_m$ will be C_m minus C_b.
All the evaluations of the regression models will be provided by eViews, in contrast to the evaluations of Neagoe (2015) which are provided by STATA.

6.4 Multinomial logit model - operational efficiency measures

After investigating the multinomial logit model we obtained time is an important and significant factor of the model. The time of the total route is related to the efficiency of the corresponding port. Shippers are interested in indirect costs associated with delays, loss of markets/market share, loss of customer confidence and opportunities forgone to inefficient service. To investigate these aspects, they propose to adjust the size of the terminal, available tugs of the port, the number of cranes and berths, but also the quality of the cranes, the quality and effectiveness of information systems and if there is an approach channel provided by the port to the model (Tongzon, 2009). To implement variables related to efficiency of a port we suggest operational efficiency measures such as capital and labor productivity, and asset utilization rates (Tongzon and Ganesalingam, 1994). These variables are, as mentioned in section 5, the number of TEUs divided by the number of employees (*employee rate*), the number of TEUs divided by the number of available tugs of the port (*tug rate*), the number of TEUs divided by the length of the berth of the port (*berth rate*) and the number of TEUs divided by the size of the terminal area (*terminal rate*).

$$\begin{aligned} \ln\left(\frac{\mathrm{P}_m}{\mathrm{P}_b}\right) &= \alpha + \beta_1 \times \Delta\mathrm{C}_m + \beta_2 \times \Delta\mathrm{T}_m + \beta_3 \times \Delta\mathrm{M}_m + \beta_4 \times \mathrm{D}_{rail} + \beta_5 \times \mathrm{D}_{barge} \\ &+ \beta_6 \times \mathrm{D}_{same\,country\,shipment} + \beta_7 \times \mathrm{D}_{ownership(shipping\,line)} \\ &+ \beta_8 \times \mathrm{D}_{ownership(terminal\,op)} + \beta_9 \times \Delta\mathrm{Employee\,rate} + \beta_{10} \times \Delta\mathrm{Terminal\,rate} \\ &+ \beta_{11} \times \Delta\mathrm{Tug\,rate} + \beta_{12} \times \Delta\mathrm{Berth\,rate} + \varepsilon. \end{aligned} \tag{23}$$

$$\begin{aligned} \ln\left(\frac{\mathrm{P}_m}{\mathrm{P}_b}\right) &= \alpha + \beta_1 \times \Delta\mathrm{C}_m + \beta_2 \times \Delta\mathrm{T}_m + \beta_3 \times \Delta\mathrm{M}_m + \beta_4 \times \mathrm{D}_{rail} + \beta_5 \times \mathrm{D}_{barge} \\ &+ \beta_6 \times \mathrm{D}_{same\,country\,shipment} + \beta_7 \times \mathrm{D}_{ownership(shipping\,line)} \\ &+ \beta_8 \times \mathrm{D}_{ownership(terminal\,op)} + \beta_9 \times \Delta\mathrm{Employee\,rate} + \beta_{10} \times \Delta\mathrm{Terminal\,rate} \\ &+ \beta_{11} \times \Delta\mathrm{Port\,berth} + \beta_{12} \times \Delta\mathrm{Port\,depth} + \varepsilon. \end{aligned} \tag{24}$$

$$\begin{aligned} \ln\left(\frac{\mathrm{P}_m}{\mathrm{P}_b}\right) &= \alpha + \beta_1 \times \Delta\mathrm{C}_m + \beta_2 \times \Delta\mathrm{T}_m + \beta_3 \times \Delta\mathrm{M}_m + \beta_4 \times \mathrm{D}_{rail} + \beta_5 \times \mathrm{D}_{barge} \\ &+ \beta_6 \times \mathrm{D}_{same\,country\,shipment} + \beta_7 \times \mathrm{D}_{ownership(shipping\,line)} \\ &+ \beta_8 \times \mathrm{D}_{ownership(terminal\,op)} + \beta_9 \times \Delta\mathrm{Employee\,rate} + \beta_{10} \times \Delta\mathrm{Terminal\,rate} \\ &+ \beta_{11} \times \Delta\mathrm{Port\,berth} + \varepsilon, \end{aligned} \tag{25}$$

where Δ will indicate the difference of the characteristics of specific route m compared to the characteristics of base route b. For example, $\Delta\mathrm{C}_m$ will be C_m minus C_b.
The transformations of the model are provided due to insignificant effects of variables, multicollinearity between variables and omitted variables. These transformations will be elaborated in section 7. Again all the evaluations of the regression models will be provided by eViews.

7 Results

In this section we will evaluate the results of the regressions with the dependent variables *throughput* and *transit*. We will elaborate the effects of the explanatory variables and make up our conclusions. Next we will evaluate the results of the multinomial logit model and compare these results to the findings of Neagoe (2015). Finally we will evaluate the results of the included operational efficiency measures.

7.1 Regression results with dependent variable throughput

In this subsection we will discuss regressions defined in section 6. The first regressions we will evaluate contain equation 1 up to and including equation 6. We will clarify the effects between the explanatory variables *GDP*, *trade* and *population* and the dependent variable *throughput*. We will expand to a logarithmic transformation. The reason for this transformation is less variation in the error terms which will be more close to normal distributed. Finally, we will evaluate the effect of the explanatory variables of the countries Romania, Bulgaria, Croatia, Greece and Slovenia compared to the other 19 European countries. The results of the regressions and the corresponding estimated coefficients can be obtained in table 3 in the appendix A.

The first point of attention is a relatively high r-squared of 0.5827 in equation 1. So 58.27% of the variation in the dependent variable *throughput* could be clarified by the explanatory variable *GDP*. A positive and significant coefficient of the explanatory variable *GDP* confirms our assumption of a positive effect between *throughput* and *GDP*.
In the second equation, we notice a relatively low r-squared, especially compared to equation 1. In this equation the variables are processed by the first differences. Although the coefficient of *GDP* is significant, we could not state there is an effect between the first differences of *throughput* and *GDP*.
In the following equations the variables are processed by a logarithmic transformation. Following by this transformation, we obtain in equation 3 an r-squared around 38%. This result suggests there is relatively less variation in the error terms which will be more close to normal distributed. Thus this transformation creates a lot of value to the explanatory variable. By adding the logarithmic transformations of the explanatory variables *trade* and *population*, we obtain an adjusted r-squared of 0.8251. This result states we can get a lot of value out of the variables. The adjusted R-squared penalizes the model for adding explanatory variables which do not improve the existing model.
By introducing a few binary variables which are linked to certain countries, we can state if the local economies depend on the national port compared to the other 19 European countries. Remarkable is the result that the effect of *GDP* is not significant anymore. This could be clarified in the way that *GDP* does not have a significant effect on *throughput* concerning the countries linked to the binary variables. The same argumentation can be handled for *population*. On the other hand *trade* still has a significant effect. The results of the effect of the binary variables linked to Bulgaria and Croatia are interesting. Concerning these countries the explanatory variables have a significant and negative effect to the dependent variable compared to the other 19 European countries. Looking at Greece, we observe a positive and significant effect of the explanatory variables to the dependent variable compared to the other 19 European countries. Thus stimulating trade in the form of policies done by public authorities or investments of private institutions would be relatively effective on improving throughput of containers in Greece compared to the other 19 countries. Neagoe (2015) excluded the explanatory variable *population* in his regression of equation 6. We select to include this variable because of the significant effect which can be obtained in equation 5.
A problem of equation 5 and 6 consists of heteroscedastistic error terms according to the Breusch-Pagan test. This means there is a correlation between the explanatory variables and the error terms. This results in a poor evaluation of the variance of the error terms caused by a biased estimation. The reason could be the explanatory variable *population* which is added in the fifth equation. By evaluating equation 6 without this variable, there still exists heteroscedasticity. So we hold on this variable in the evaluation. A solution for this problem could be using instrumental variables. These variables are highly correlated with the explanatory variables, but poorly correlated to the error terms.
Overall, we mostly obtain positive and significant coefficients of the explanatory variables which confirms a positive and significant effect to the dependent variable *throughput*. The only excep-

tion consists of *population* in equation 5, which refers to a negative and significant effect. By adding each of the logarithmic transformations of the explanatory variables *trade* and *population* we create more explanatory power to the regression. Both the variables are significant at a 5% significance level and adding these variables increases the adjusted r-squared, thus both of the variables are valuable. By introducing binary variables linked to the given countries, we did investigate these countries effect of the explanatory variables to the dependent variable compared to the other 19 European countries.

7.2 Regression results with dependent variable transit

In this subsection we will discuss the following seven regressions defined in section 6. The regressions we will evaluate contain equation 7 up to and including equation 12. We will clarify the effects of the explanatory variables *density*, *trade*, *GDP* and *population* to the dependent variable *transit*. We will extend to a logarithmic transformation. The reason for this transformation is less variation in the error terms which will be more close to normal distributed. Equation 12 will include clusters in the regression. The modeling of the clusters will be clarified in this subsection and the process of adding clusters to the regression is clarified in section 6. The results of the regressions and the corresponding estimated coefficients can be obtained in table 4 in the appendix B.

The linear regression of the seventh equation shows a positive and significant effect of the explanatory variable *density* to the dependent variable *transit*. This result could be clarified as an increase in transit cargo for every extra meter of highway density. To improve the relatively low fit of the model, a logarithmic transformation is introduced in equation eight. The r-squared of this regression corresponds to 0.3618. So 36.18% of the variation in the dependent variable *transit* could be clarified by the logarithmic transformation of the explanatory variable *density*. We can not compare the r-squared between equation 7 and 8 because of the change in the variance of the explanatory variables caused by the logarithmic transformation.
By adding the logarithmic transformation of the explanatory variables *trade*, *GDP* and *population* we obtain an r-squared of 0.5939 in equation 11. There is still a gap around 40% of the variation in the dependent variable *transit* which can not be clarified by the explanatory variables, so we have to be careful with the statements concerning this regression.
The point which strikes us is a negative significant effect of the explanatory variable *GDP* to the dependent variable *transit* in equations 10 and 11. When acquiring the correlations we notice a correlation of 0.97 between *GDP* and *trade*. This could be a causality for the negative significant effect of *GDP*.

We need to generate clusters to show the difference of the effect from the explanatory variables to the dependent variable between small, medium and large economies. A cluster contains a set of countries of Europe ordered based on similar statistics of *density*, *trade*, *GDP* and *population*. Cluster 2 contains countries with the smallest economies of Europe, cluster 3 contains countries with the largest economies of Europe and cluster 1 contains the rest of countries of Europe. To divide the countries over the clusters we follow the sets of Neagoe (2015). The only problem is Slovakia, which is divided over cluster 1 and cluster 2. Before allocating Slovakia to a cluster, we will state the importance of the explanatory variables according to the significance of the explanatory variables in the equations and the adjusted r-squared. In all the equations all the explanatory variables are significant at a 5% significance level, thus we will focus on the adjusted r-squared. After acquiring the adjusted r-squared of equation 8, 9, 10 and 11, the adjusted r-squared equals around 99% of the corresponding r-squared. These results state that every explanatory variable is equally important. We analyze the different statistics of cluster 1 and cluster 2, because the statistics of cluster 3 are out of range and therefore not relevant to include in this process. Slovakia's explanatory variable *density* is most corresponding to cluster 1 and explanatory variables *GDP*, *population* and *trade* are most corresponding to cluster 2, thus we add Slovakia to cluster 2. These statistics can be obtained in table 5 in the appendix B.

After adding the clusters to the model of equation 12 and creating the model of equation 13, we obtain the results of the regressions and the corresponding estimated coefficients in table 6 in the appendix B. We obtain a positive and significant coefficient according to the dependent variable when including the first and second cluster separately. Adding cluster 3 in the regression appears to have a negative and significant coefficient. Including cluster 1 and cluster 2 together in the model gives a positive significant coefficient for both of the clusters as well. On top of this result, the r-squared increases to 0.6283 which corresponds to a better fit of the model. The adjusted r-squared increases to 0.6178 which corresponds to more explanatory power of the explanatory variables. This result could be clarified as a positive effect of improving the infrastructure of relatively small economies to increase the amount of transit cargo through these economies, where the relatively big economies will not benefit of these improvements.

7.3 Results multinomial logit model

In this subsection we will evaluate the results of the multinomial logit model which includes the equations 17 up to and including equation 22. We will clarify the effects of the explanatory variables *transport cost*, *transport time*, *crane congestion*, terminal handling costs (THC), *port depth* and *port berth* according to the relative market shares of route m compared to the relatives market share of the base route. These routes focus on the transport from one of the eight ports to the hinterland regions of one of the eight countries, both mentioned in section 5. Furthermore, we will introduce binary variables to link the explanatory variables to characteristics of the transport lane. We add these variables and binary variables to the model cause they potentially influence the port choice of a shipper Neagoe (2015). All the equations contain 850 observations which are included according to the route route selection procedure described in section 5. The results of the estimated coefficients can be obtained in table 7 in the appendix C.

The first point which strikes us relates to the exclusion of the constant in equation 17. When the explanatory variables force the regression line to go through the origin, it would confirm the only possibility to exclude the constant to the regression. Otherwise, the variables will be estimated with biased coefficients, even if the constant is not significant. Since the explanatory variables are processed with the difference between route m and the base route, this could be possible. For example, we estimate a function for the *transport costs* for every type of transport according to a regression of the *transport distance* to the *transport costs* as described in section 5. This function includes a constant, but will be eliminated after subtraction. For the other explanatory variables we could apply the same reasoning. We observe a negative r-squared for equation 17, which means that the fit of the model is actually worse than just fitting a horizontal line. In this case the r-squared cannot be interpreted as the square of a correlation. Such situations indicate that a constant term should be added to the model.
When we add a constant to the model (18), we observe an r-squared of 0.1253. This implies still a relatively low fit of the model, but it proves the model needs a constant to interpret the results of the explanatory variables. We observe a negative and significant coefficient at a 1% significance level for the *transport time*, which implies that extra time of the selected route would have a negative effect on the utility of the shipper. We observe a negative and significant coefficient at a 5% level for the *maritime diversion distance*, which implies that an extra nautical mile out of the main trading lane would have a negative effect for the utility of the shipper as well. These two observations correspond to our intuition. Although the positive coefficient of the *transport cost* is remarkable. This would imply that an extra euro spend on the transport costs would

have a positive effect on the utility of the shipper. However this coefficient is insignificant, thus we cannot confirm the effect of this statement.
Adding the binary variables related to rail and barge transport to the model increases the fit of the explanatory variables of the model from an adjusted r-squared of 0.1222 to 0.1797. This appearance clarifies we get more value out of the explanatory variables compared to the previous model. The sign of the coefficients of the *transport cost*, *transport time* and *maritime diversion distance* stay the same, but the coefficient of the *maritime diversion distance* is now even significant at a 1% level. We observe a positive and significant coefficient for rail transport, which implies the shipper would prefer rail transport to the hinterland regions over road transport. This could be related to the negative and significant coefficient of the *transport time* for the selected route.
Equation 20 includes the explanatory variables of the previous models and the binary vector corresponding to transport within the same country as well. We observe a positive and significant coefficient for the binary variable, which implies a preference of the shipper for a route within the same country. This evaluation corresponds to our intuition, because transport over the border could impose restrictions or other implications.
The next model adds the binary variables corresponding to the ownership and management of a terminal. Three alternatives are possible: the terminal is owned and managed by publicly port authorities, the terminal is owned and managed by terminal operators or the terminal is owned and managed by terminal operators affiliated to shipping lines. We observe a negative and significant coefficient at a 1% level for the binary variables corresponding to the situations when the terminal is owned and managed by terminal operators, and when the terminal is owned and managed by terminal operators affiliated to shipping lines. These results imply a preference of the shippers for ports which contain terminal owned and managed by publicly port authorities. Mostly this form of ownership and management would increase the regulations and processes are often inefficient compared to outsourcing to the private sector, thus our intuition would disagree with this alternative. On the other hand, there could be more attention for safety.
The last model includes variables which can be seen as characteristics corresponding to a specific port. The included variables are *crane congestion*, *THC* (terminal handling costs), *port berth* and *port depth*. The first point of interest is the change of sign of the *maritime diversion distance*, although the coefficient is insignificant. Another interesting aspect is the change of sign of the binary variable corresponding to the ports when the terminal is owned and managed by terminal operators affiliated to shipping lines. This coefficient is insignificant, so we cannot conclude a preference of a shipper to this effect. After investigating the correlation between the explanatory variables we observe a correlation of 0.9340 between *port depth* and the binary variable corresponding to the terminals owned and managed by terminal operators affiliated to shipping lines. The high correlation could be caused by restrictions the corresponding management applies to the qualities of the port. In this case, they could be very restrictive to the port depth. This probably influences the change of sign. Note that we cannot state the preference of a shipper in the previous model as well. However we can state that the terminals owned and managed by terminal operators are less preferred than the terminal owned and managed by publicly port authorities. Furthermore we observe that the added variables all have a negative effect to the choice of a specific route, except for the explanatory variable *port depth*. On top of the positive effect, *port depth* is the only variable which has a significant effect at a 1% level. Every extra meter of water depth would increase the probability of a shipper to select a corresponding route according to the model. This final model has an r-squared of 0.4032 and a relatively high corresponding adjusted r-squared of 0.3946. So 40.32% of the variation in the choice of a shipper for a selected route could be clarified by the explanatory variables.
Overall we observe significant differences with the multinomial logit model of Neagoe (2015). Before we analyze the differences in the results we will evaluate the possible differences in the data used in the models. As described in section 5 we estimated the *transport costs* according

to a function based on the regression of the transport distance to the *transport costs*. We made an estimation for the *port berth* and *port depth* as well. The amount of gantry cranes available in Burgas is gathered from port of Burgas (2017). Now we will evaluate the differences between our model and the model of Neagoe (2015). The first point of attention relates to the excluded constant in the model of Neagoe (2015). As mentioned before when the explanatory variables force the regression line to go through the origin, it would confirm the only possibility to exclude the constant to the regression. Probably this is not the case in our multinomial logit model. Next there is an opposite effect observed in both models for the explanatory variable *transport time*. The result of our model is more logical and confirms our intuition. The opposite effect of *transport cost* is striking as well, but the positive effect in our model is not significant so we cannot state a conclusion to this result. According to Neagoe (2015) shippers would prefer a route by road transport from the port to the hinterland regions. Interpretations of our model clarifies the shipper would prefer rail transport. This difference could be related to the difference in the routes selected for our model and the model of Neagoe (2015). As clarified in section 5 we suggested a preference for road transport for the relatively shorter routes and rail transport for relatively longer routes. Neagoe (2015) observed a preference of the shipper for terminals owned and managed by terminal operators compared to terminals owned and managed by publicly port authorities. We observe the opposite effect. Finally we observe an opposite effect of the explanatory variable *port depth*. Our intuition would state that an increase in the water depth would increase the opportunities for the shipper, but could increase the restrictions for the cargo ships as well because of the probabilities of relatively larger cargo ships which could enter the port.

7.4 Results multinomial logit model - operational efficiency measures

In this subsection we will evaluate the results of the multinomial logit model - operational efficiency measures which include the equations 23 up to and including equation 25. We will clarify the interesting effects of the explanatory variables of the original multinomial logit model, and *employee rate*, *terminal rate*, *tug rate* and *berth rate* according to the corresponding port of route m compared to the port of Constanza. These extra variables could potentially influence the port choice of a shipper (Tongzon and Ganesalingam, 1994). The results of the estimated coefficients can be obtained in table 8 in the appendix D.

After adding the operational efficiency measures to the model of equation 21 we obtain mostly the same effects for the nested explanatory variables of equation 21. An interesting difference is the change of sign of the coefficient related to the binary variable when the terminal is owned and managed by terminal operators affiliated to shipping lines, and the negative and significant coefficient related to the binary variable when the terminal is owned and managed by terminal operators. This results states the preference of a shipper to the terminal owned and managed by terminal operators affiliated to shipping lines. All of the added explanatory variables contain a significant coefficient at a 1% level to the dependent variable. First, we observe a positive coefficient for the *employee rate* and the *tug rate* which could be clarified by an increasing ratio of number of TEUs handled in the port to the employees in the port and an increasing ratio of number of TEUs handled in the port to the available tugs of the port would increase the probability of a shipper to select this port. Thus high capital and labor productivity would increase the probability a shipper would select this route m with the corresponding port. Then a negative coefficient for the *terminal rate* and *berth rate* could be clarified by an increasing ratio of number of TEUs handled in the port to the size of the terminal of the port and an increasing ratio of number of TEUs handled in the port to length of the berth would decrease the probability of a shipper to select this port. Thus, increasing asset utilization rates decrease the probability of a shipper to select a specific route corresponding to this port. A reason for

this result could be an aversion for congestion which could lead to unsafe situations in the port. After inspecting the correlation between the explanatory variables we obtain a relatively high correlation of 0.9772 between *tug rate* and *berth rate*. When including only one of these variables to the model we obtain an insignificant effect for both of the variables which indicates to multicollinearity between these variables. Thus we eliminate these variables from the model. We add *port berth* to the model to keep the asset utilization related variable in the model and add *port depth* to the model, because of the significant coefficient in equation 22. The points which strike us are the insignificant coefficients of *employee rate*, *terminal rate*, *port berth* and *port depth*. After inspecting we do not obtain abnormal high correlations between these variables which could cause these results (table 9 in the appendix D).
If we eliminate *port depth* we obtain an increase of the adjusted r-squared which clarifies we get more value out of the explanatory variables than the previous model. Thus *port depth* can be seen as an omitted variable which contains information irrelevant to the model, so we eliminate this variable from the model. The first interesting aspect of equation 25 is the positive and significant coefficient of *maritime diversion distance* (M_m). This result could be clarified by an increase of the nautical miles of the port from the main trading line would increase the probability of the shipper to select this route. This result is contrary to our intuition which disagrees for a preference of an inefficient route. This result could be caused by a relatively high negative correlation to the *employee rate* (-0.8798) and *terminal rate* (-0.8267). If we eliminate the *maritime diversion distance* from the model we obtain a significant drop of the r-squared to 0.3754 and adjusted r-squared to 0.3680, so we do not prefer this procedure. Finally, we obtain significant coefficients for *employee rate* and *terminal rate* which enables us to state the same preference for the shipper as we did in the results of equation 23.
Overall, we prefer the last model because of the increased adjusted r-squared which relates to the improving power of the explanatory variables. Our most important findings are the preference of a shipper for terminals owned and managed by terminal operators affiliated to shipping lines, a preference for a relatively high level of capital and labor productivity, and a preference for a relatively low level of asset utilization.

8 Conclusion and further research

Inspecting the data of the possible routes yields interesting findings. When focusing on time efficiency, based on average speed and corresponding distance, we discovered transport within 700 kilometers would be most efficient with road transport, and transport over 700 kilometers would be most efficient with rail transport. According to Neagoe (2015) we suggested rail transport to be cheaper than road and barge transport on every route. We obtain this is not always the case and has to be considered in every situation.
Then we evaluated regressions concerning the dependent variable *throughput*. We observed positive and significant coefficients of the explanatory variables *GDP* and *trade*, and a negative and significant coefficient of the explanatory variable *population*. Broadly speaking, stimulating the economy by imports and exports, and governmental policies would increase the throughput in a country, which could improve the economy of Romania. After including binary variables corresponding to the countries in South-East Europe, we could test the effect of these countries to the dependent variable compared to the other 19 European countries. We observed a negative and significant coefficients of the binary variables related to Bulgaria and Croatia, and a positive and significant coefficient of the binary variables related to Greece. We concluded that stimulating imports and exports done by public authorities or private institutions would be relatively effective on the economical development in Greece compared to the other 19 countries. We cannot state an expected effect corresponding to Romania because of the insignificant coefficient.
Next we investigated the effects of *density*, *trade*, *GDP* and *population* to the amount of cargo

transiting a country. We obtained that increasing the highway density in a country and stimulating imports and exports would increase the amount of cargo transiting a country resulting in a boost for the economy. An interesting point is the negative and significant coefficient of *GDP* to *transit*. *GDP* appears to be highly correlated to *trade* which could create this effect. Clusters, including countries of Europe with different economical sizes, would show the effects of the explanatory variables in relatively small and relatively big economies. The main result of these clusters could be clarified as a positive effect of improving the infrastructure of relatively small economies to increase the amount of transit cargo through these economies, where the relatively big economies will not benefit from these improvements.

Then we investigated the multinomial logit model with the relative market share of selected route m compared to the base market share of the base route as the dependent variable. We firstly obtained the effects of the explanatory variables *transport cost*, *transport time* and *maritime diversion distance*. *Transport time* and *maritime diversion distance* provided the expected negative and significant coefficient which means that every extra minute of transport time and every extra nautical mile the concerning port is deviating of the main trading lane would decrease the preference level of the shipper to select this corresponding route. The following interesting result is the preference of shippers of rail transport over road transport for transport from the port to the hinterland regions. This would increase the motivation of improving the connection between ports and railway stations. Furthermore we observed a preference of shippers for routes within the same country. In the following expanded model we obtained a preference of the shipper for terminals owned and managed by publicly port authorities. This result is remarkable in the sense that mostly this form of ownership would increase the regulations to the corresponding port. The final interesting result is an extra meter of water depth in the port would increase the preference of the shipper for this port.

After investigating the multinomial logit model we obtained time to be an important and significant factor of the model. A decline in the time spend per route would increase the probability of a shipper to select this specific route. Because of the fact the time of the route is related to the efficiency of the port we suggested to include operational efficiency measures to the model. These measures contain capital and labor productivity, and asset utilization rates. Our most important findings are the preference of a shipper for a relatively high level of capital and labor productivity, and a preference for a relatively low level of asset utilization. We advice to keep these ratios in mind to increase the attractiveness of the port for the shippers and thereby improve the economy of the country.

In the regressions related to the dependent variable *throughput* we obtained heteroscedasticity in equations 5 and 6. This result followed after including the explanatory variable *population* to the regression. To solve this problem we could introduce instrumental variables which are highly correlated to *population* but not correlated to the error terms of the regression. Further research could be focused on selecting these instrumental variables. In the regressions related to the dependent variable *transit* we observed high correlation of 0.97 between the explanatory variables *GDP* and *trade*. This correlation could result in misleading effects of these variables. In this case instrumental variables for these variables could be considered as well. According to Slack (1985) mostly shippers are conservative in their port choice. Further research could expand the multinomial logit model to include proxy variables related to this aspect to evaluate this effect. In our multinomial logit model we are focused on goods and products. In our opinion it would be interesting to evaluate the effects of liquids and fuels in the model. Then data of transport through pipelines could be included as well. According to Tongzon (2009) cargo safety is of great importance for transport, sometimes even more important than the actual safety. Customers appreciate quality and reliability of the delivery of their products. A bad reputation of the port could scare off potential customers and discourage current customers. Further research could include the effect of cargo accidents in the model.

References

Bird, J. and Bland, G. (1988). Freight forwarders speak: the perception of route competition via seaports in the european communities reasearch project. *Maritime Policy & Management*, 15(1):35–55.

Cazzinga, F., Foschi, D., and Foschi, A. (2002). Mediterranean versus northern range ports. why do italian containers still prefer routing via the northern range ports? advice for a new policy. *IAME 2002 - Panama*, 1(2):1–19.

container insight (2014-1-25). http://ciw.drewry.co.uk/features/mediterranean-transhipment-booms/#.vmtrhf7f8xs.

de Langen, P. (2007). Port competition and selection in contestable hinterlands: the case of austria. *Department of Port, Transport and Regional economics*, 7(1):1–14.

ETISplus (2014). http://ftp.demis.nl/outgoing/etisplus/datadeliverables/.

eurostat (2018). http://ec.europa.eu/eurostat/data/database.

Ferrari, C., Parola, F., and Morchio, E. (2006). Southern european ports and the spatial distribution of edcs. *Maritime Economics & Logistics*, 8(1):60–81.

global agent (2018). http://www.globalagent.hr/ports/port-of-rijeka-359.html.

harbours review (2018a). http://harboursreview.com/port-thessaloniki.html.

harbours review (2018b). http://harboursreview.com/rijeka-port.html.

Keeler, T. and Ying, J. (1988). measuring the benefits of a large public investment: The case of the us federal-aid higway system. *Journal of Public Economics*, 36(1):69–85.

Malchow, M. (2004). A disaggregate analysis of port selection. *Transportation Research Part E: Logistics and Transportation Review*, 40(4):317–337.

McKinnon, A. (2007). Decoupling of road freight transport and economic growth trends in the uk: an exploratory analysis. *Transport Reviews*, 27(1):37–64.

medcruise (2018a). http://www.medcruise.com/port/526/thessaloniki/information.

medcruise (2018b). http://www.medcruise.com/port/702/constantza/information.

Medda, F. and Carbonaro, G. (2007). Growth of container seaborne traffic in the mediterranean basin: Outlook and policy implications for port development. *Transport reviews*, 27(5):573–587.

Neagoe, M. (2015). Constanza-container gateway. *Urban, Port and Transport Economics*, June(1):1–96.

Notteboom, T. (2010). Concentration and the formation of multi-port gateway regions in the european container port system: an update. *Journal of Transport Geography*, 18(4):567–583.

Notteboom, T. (2013). Recent traffic dynamics in the european container port system. *Port Technology International Issue*, 58(1358-1759):14–18.

Notteboom, T. and Dooms, M. (2014). State of the european port system - market trends and structure update. *Partim transsphipment volumes*, 1(1):1–31.

OECD (2018). https://data.oecd.org/emp/employment-rate.htm.

port of Burgas (2017). http://navbul-portburgas.com/wp-content/uploads/presentations/bmf_port_burgas_2017.pdf.

port of Constanza (2018). http://www.portofconstantza.com/apmc/i.do?lan=en.

port of Koper (2018). https://www.luka-kp.si/eng/information-for-ships.

port of Napa (2018a). http://www.portsofnapa.com/port-of-koper.

port of Napa (2018b). http://www.portsofnapa.com/port-of-trieste.

port of Piraeus (2018). http://www.worldportsource.com.

port of Rijeka (2018). http://www.lukarijeka.hr.

port of Trieste (2018). http://www.porto.trieste.it/eng/port/technical-information.

port of trieste (2018). http://www.trieste-marine-terminal.com/en/personnel-and-working-time.

port of Varna (2018). http://port-varna.bg/bg/terminals/varna-iztok.

portfleet99 (2018). http://www.portfleet99.com.

Rodrique, J., Comtois, C., and Slack, B. (2013). *The Geography of Transport Systems.* Routledge, Third Avenue, New York.

Rodrique, J. and Notteboom, T. (2010). foreland-based regionalization: Integrating intermediate hubs with port hinterlands. *Research in Transport Economics*, 27(1):19–29.

sea distances (2018). https://sea-distances.org/.

Slack, B. (1985). Containerization, inter-port competition, and port selection. *Maritime Policy & Management*, 12(4):293–303.

Tongzon, J. (2009). port choice and freight forwarders. *Transportation Research Part E: Logistics and Transportation Review*, 45(1):186–195.

Tongzon, J. and Ganesalingam, S. (1994). An evaluation of asean port performance and efficiency. *Asian economic journal*, 8(3):317–330.

tradingeconomics (2018). https://tradingeconomics.com/croatia/employment-rate.

Tsamourgelis, I., Paflioti, P., and Vitsounis, T. (2013). Seaports activity (a)synchronicity, trade intensity and business cycle convergence: Apanel data analysis. *International Journal of Maritime , Trade & Economic Issues*, 1(1):67–92.

van Hooydonk, E. (2013). http://www.ericvanhooydonk.be/media/2013-01-08-ec-port-labour-study-vol2.pdf.

Veldman, S. and Buckmann, E. (2003). a model on container port competition:an application for the west european container hub-ports. *Maritime Economics & Logistics*, 5(1):3–22.

Appendices

A Section 7.1

Table 3: Regression results with dependent variable throughput

Coefficients	Equation 1	Equation 2	Equation 3	Equation 4	Equation 5	Equation 6
Constant	651.47	0.04	-3.42	-3.93	-2.89	-3.73
T-statistic	4.26**	2.68**	-11.89**	-13.49**	-5.33**	-6.74**
GDP	0.00	1.09	0.85	0.31	0.44	0.08
T-statistic	20.74**	6.82**	35.71**	3.08**	3.80**	0.69
Trade				0.59	0.58	0.75
T-statistic				5.34**	5.28**	6.75**
Population					-0.15	0.05
T-statistic					-2.27*	0.66
D_{Rom}						-0.22
T-statistic						-0.99
D_{Bul}						-1.08
T-statistic						-5.13**
D_{Cro}						-1.02
T-statistic						-4.73**
D_{Gre}						0.68
T-statistic						3.76**
D_{Slo}						-0.26
T-statistic						-1.41
R-squared	0.5827	0.1407	0.8055	0.8238	0.8268	0.8571
Adj. r-squared	0.5814	0.1377	0.8049	0.8227	0.8251	0.8532
Observations	310	286	310	308	308	308

*, ** denote the coefficients are significant at a 5%, 1% significance level respectively. In equation 2, the variables are processed by the first differences. Adj. r-squared defines the adjusted r-squared. From equation 3 up to and including equation 6, the dependent variable and explanatory variables are processed by a logarithmic transformation. D_{Rom}, D_{Bul}, D_{Cro}, D_{Gre} and D_{Slo} denote Romania, Bulgaria, Croatia, Greece and Slovenia, respectively. If an observation refers to one of these countries, it will contain the value 1, otherwise 0. The equations are visible in section 6.

B Section 7.2

Table 4: Regression results with dependent variable transit

Coefficients	Equation 7	Equation 8	Equation 9	Equation 10	Equation 11
Constant	6169.46	8.47	4.15	1.80	-0.49
T-statistic	3.49**	95.16**	5.12**	2.36*	-0.47
Density	4311.73	0.84	0.67	0.58	0.73
T-statistic	6.67**	11.12**	8.55**	8.39**	8.76**
Trade			0.37	2.29	2.19
T-statistic			5.35**	9.70**	9.34**
GDP				-1.69	-1.95
T-statistic				-8.42**	-9.12**
Population					0.41
T-statistic					3.11**
R-squared	0.1693	0.3618	0.4363	0.5757	0.5939
Adj. r-squared	0.1655	0.3589	0.4311	0.5698	0.5863
Observations	220	220	220	220	220

*, ** denotes the coefficients are significant at a 5%, 1% significance level respectively. Adj. r-squared defines the adjusted r-squared. From equation 8 up to and including equation 11, the dependent variable and explanatory variables are processed by a logarithmic transformation. The equations are visible in section 6.

Table 5: Economic statistics

		Density	GDP	Population	Trade
Slovakia	average	0.75	53’446	5’381’118	10’510
	std. deviation	0.09	15’539	11’085	2’035
Cluster 1	average	0.95	207’441	13’332’827	156’427
	std. deviation	0.86	227’953	12’509’284	126’873
Cluster 2	average	1.76	68’440	8’171’350	53’828
	std. deviation	1.96	99’087	12’552’309	70’041

Cluster 1 and cluster 2 contain countries of Europe ordered based on similar statistics of transit, density, trade, GDP and population. Std. deviation defines the standard deviation.

Table 6: Regression results with variation in clusters, equations 12 and 13

Coefficients	12.1	12.2	12.3	13
Constant	-0.79	-0.80	-4.41	-8.50
T-statistic	-0.75	-0.55	-3.13**	-3.80**
Ln(Density)	0.85	0.71	0.94	0.89
T-statistic	8.36**	7.55**	9.75**	9.00**
Ln(Trade)	2.09	2.21	2.18	2.32
T-statistic	8.79**	8.96**	9.62**	9.76**
Ln(GDP)	-1.97	-1.93	-1.86	-1.76
T-statistic	-9.29**	-8.90**	-8.95**	-8.32**
Ln(Population)	0.51	0.40	0.61	0.60
T-statistic	3.66**	2.97**	4.47**	4.36**
Cluster 1	0.38			1.42
T-statistic	2.09*			4.43**
Cluster 2		0.09		1.98
T-statistic		0.31**		3.88**
Cluster 3			-1.20	
T-statistic			-4.00**	
R-squared	0.6020	0.5941	0.6221	0.6283
Adj. r-squared	0.5927	0.5846	0.6133	0.6178
Observations	220	220	220	220

*, ** denote the coefficients are significant at a 5%, 1% significance level respectively. Adj. r-squared defines the adjusted r-squared. Cluster 1, cluster 2 and cluster 3 contain countries of Europe ordered based on similar statistics of transit, density, trade, GDP and population. 12.1, 12.2 and 12.3 denote equation 12 including cluster 1, cluster 2 and cluster 3 respectively. 13 denotes equation 13. The equations are visible in section 6.

C Section 7.3

Table 7: Results multinomial logit model

Coefficients	Eq. 17	Eq. 18	Eq. 19	Eq. 20	Eq. 21	Eq. 22
Constant		1488	1130	557	1034	3164
T-statistic		13.39**	9.24**	4.47**	7.48**	5.85**
Δ C_m	0.0012	0.0001	0.0003	0.0006	0.0002	0.0008
T-statistic	0.56	0.06	0.15	0.34	0.12	0.51
Δ T_m	-1.1018	-0.9247	-0.8423	-0.7026	-0.5668	-0.4791
T-statistic	-11.80**	-10.77**	-10.06**	-8.91**	-7.19**	-6.21**
Δ M_m	-2.9048	-0.7006	-0.8454	-1.3798	-2.2423	0.0631
T-statistic	-9.10**	-2.10*	-2.60**	-4.51**	-4.91**	0.08
D_{rail}			1757	1266	1199	1074
T-statistic			7.40**	5.62**	5.55**	5.14**
D_{barge}			-1129	-799	381	1074
T-statistic			-1.86	-1.42	0.68	1.88
D_{same}				2556	2498	2871
T-statistic				11.42**	11.41**	13.20**
$D_{shipping\,line}$					-1988	-1890
T-statistic					-3.74**	-1.64
$D_{terminal\,op}$					-1651	-4510
T-statistic					-8.36**	-8.70**
Δ Crane						-0.0108
T-statistic						-1.35
Δ THC						-14.5062
T-statistic						-1.84
Δ Port berth						-0.2857
T-statistic						-0.89
Δ Port depth						643
T-statistic						6.76**
R-squared	-0.0601	0.1253	0.1846	0.2938	0.3524	0.4032
Adj. r-squared	-0.0626	0.1222	0.1797	0.2888	0.3463	0.3946
observations	850	850	850	850	850	850

*, ** denote the coefficients are significant at a 5%, 1% significance level respectively. Adj. r-squared defines the adjusted r-squared. D_{same}, $D_{shipping\,line}$, $D_{terminal\,op}$ and Δ Crane is an abbreviation for $D_{same\,country\,shipment}$, $D_{ownership(shipping\,line)}$, $D_{ownership(terminal\,op)}$ and Δ Crane congestion respectively. Eq. is an abbreviation for equation and the equations are visible in section 6.

D Section 7.4

Table 8: Results multinomial logit model - time efficient

Coefficients	Equation 23	Equation 24	Equation 25
Constant	-148	2449	1865
T-statistic	-0.53	2.63**	9.58**
Δ C_m	0.0008	0.0008	0.0009
T-statistic	0.51	0.51	0.52
Δ T_m	-0.4791	-0.4791	-0.4756
T-statistic	-6.21**	-6.21**	-6.18**
Δ M_m	-0.4289	5.6952	11.2745
T-statistic	-0.30	0.64	6.21**
D_{rail}	1074	1074	1066
T-statistic	5.14**	5.14**	5.11**
D_{barge}	1074	1074	987
T-statistic	1.88	1.88	1.78
D_{same}	2871	2871	2868
T-statistic	13.20**	13.20**	13.20**
$D_{shipping\,line}$	23706	6208	11611
T-statistic	6.18**	0.69	3.82**
$D_{terminal\,op}$	-1198	-3795	-3124
T-statistic	-5.05**	-3.52**	-12.07**
Δ Employee rate	1.2876	2.4217	4.6542
T-statistic	2.85**	0.68	7.63**
Δ Terminal rate	-1.1591	-0.5311	-0.9178
T-statistic	-6.91**	-0.86	-6.27**
Δ Tug rate	0.3143		
T-statistic	7.60**		
Δ Berth rate	-20.8147		
T-statistic	-7.56**		
Δ Port berth		1.3823	2.9223
T-statistic		0.57	7.59**
Δ Port depth		301	
T-statistic		0.64	
R-squared	0.4032	0.4032	0.4029
Adj. r-squared	0.3946	0.3946	0.3950
Observations	850	850	850

*, ** denote the coefficients are significant at a 5%, 1% significance level respectively. Adj. r-squared defines the adjusted r-squared. D_{same}, $D_{shipping\,line}$ and $D_{terminal\,op}$ is an abbreviation for $D_{same\,country\,shipment}$, $D_{ownership(shipping\,line)}$ and $D_{ownership(terminal\,op)}$ respectively. The equations are visible in section 6.

Table 9: Correlations between relevant variables equation 22

Correlation	Δ Employee rate	Δ Terminal rate	Δ Port berth	Δ Port depth
Δ Employee rate	1.00	0.74	0.46	0.58
Δ Terminal rate	0.74	1.00	0.88	0.40
Δ Port berth	0.46	0.88	1.00	0.38
Δ Port depth	0.58	0.40	0.38	1.00