Erasmus University Rotterdam

Econometrics and Management Science

Master Thesis

Intermodal transport in Europe: Trends & Drivers

Author:
Jurriaan Wesselink

Supervisor:
Prof. Dr. Richard Paap

Co-reader:
Dr. Wendun Wang

External supervisor:
Mitchell van Balen

20-8-2017

Abstract

European goverments aim at intermodal freight transport instead of road freight transport. An increase of intermodal freight transport could lead to an improvement of cost effectiveness, economic growth and the reduction of social and environmental externalities. Despite this, road transport still remains dominant. Quantification studies could lead to a better understanding of how to promote intermodal transport. Not a lot of econometric research has been done yet, especially because data are on country level. In this research data on a more disaggregated level, namely the level of origin-destinations, could be used. Linear, Poisson and negative binomial fixed and random effects models are implemented to declare the number of departures per week from a certain origin to a certain destination. For sake of comparison market share models are applied on country level data to describe the positions of several countries in the freight market. The results address useful and interesting relationships and show that government policy should focus on investments in port infrastructure and on increasing Diesel tax. Furthermore, it shows the importance of quantification and the importance of data on the level of origin-destination.

Keywords: Intermodal transport, Fixed effects model, Random effects model, Count models, Market share model

Acknowledgement

I will briefly express my gratitude to the people and instances that supported me during this process. I would like to thank Ecorys for giving me the opportunity to work on this project. Especially I would like to thank Mitchell van Balen for his help, support and critical notes on the qualitative part of the research. I would like to thank Jeroen Bozuwa for his support and the collection of the data. Without the data I would not have come up with new insights about intermodal transport. I would like to thank my supervisor Prof. Dr. Richard Paap for his support and guidance from the beginning until the end. On this journey his help has been clear and constructive. In addition, I would like to thank Dr. Wendun Wang for his help as co-reader.

Contents

1 Introduction

In our globalized world, the demand for freight transport is continuously rising. Most freight is still transported by road, but since the late seventies the concept of intermodality gains traction. Briefly stated, this is transport by sea, inland water or rail. Compared to transport by road, intermodal transport could lead to an improvement of cost effectiveness, economic growth and the reduction of social and environmental externalities. Despite that these advantages are recognized and supported by governments, road transport remains dominant. That is why it is so important to better understand how intermodal transport can be promoted. Quantification studies are asked for. In this research we will quantify European intermodal freight transport flows. We will investigate the following main question: "How was European intermodal freight transport organised in the past years and what factors promote a modal shift?"

Empirical models are used to unveil the intermodal freight transport flows. Not a lot of econometric research has been done, because data about freight transport flows are only available aggregated on country level. In this research we use unique and disaggregated data. We use panel data on origin-destination level from 2014 to 2017 for freight transport by rail, inland water and sea. The dependent variable is the number of departures per week from a certain origin to a certain destination. We want to declare the number of departures by (possibly) related variables as time, GDP and port quality. Therefore, linear random and fixed effects models will be used. These methods estimate parameter values for (possibly) related variables and in addition estimate the unobserved individual heterogeneity for the dependent variable. Since this concerns a count variable, we will also apply Poisson and negative binomial fixed and random effects models. Furthermore, we want to estimate unobserved heterogeneity of certain countries and TEN-T corridors (main transport routes in Europe). Therefore clustered linear and Poisson fixed and random effects models will be used. Interesting insights about European freight transport flows are generated. In addition, we apply market share models of country freight shares for road, rail and inland water transport. For this we use data aggregated on country level.

The paper is set up as follows. First we present a literature review and some further specified research questions. In the second section we address the preparation and arrangement processes and statistics of the data set(s). Then we present and explain the used models, methods and statistics. Next, we will present the results and shortly analyze these in terms of added value to the research question. Finally, we conclude and we present a discussion.

2 Literature review and research questions

Since 1970 European freight transport has increased significantly. Although governments aim to increase intermodal transport instead of road transport, the share of road transport increased compared to the share of intermodal freight transport. Research in this field is therefore important. Since the 1990s the amount of research on intermodal transport in freight distribution has grown significantly (Agamez-Arias and Moyano-Fuentes, 2017). Bontekoning, Macharis and Trip (2003) point out that intermodal transport is addressed seperately from unimodal transport in the literature, where a significantly amount of analytical publications specifically address intermodal transport. In practice it is considered as a competing mode and can be used as an alternative to unimodal transport. In addition, intermodalism has become an important policy issue.

This section will address the definition of intermodal transport, the evolution and current statistics of European intermodal transport, European policy directives and its effectiveness and some models that currently address the intermodal landscape. Furthermore, in combination with the literature and in contribution to the main question and stated problems we formulate some more concrete questions to consolidate the directives of this research.

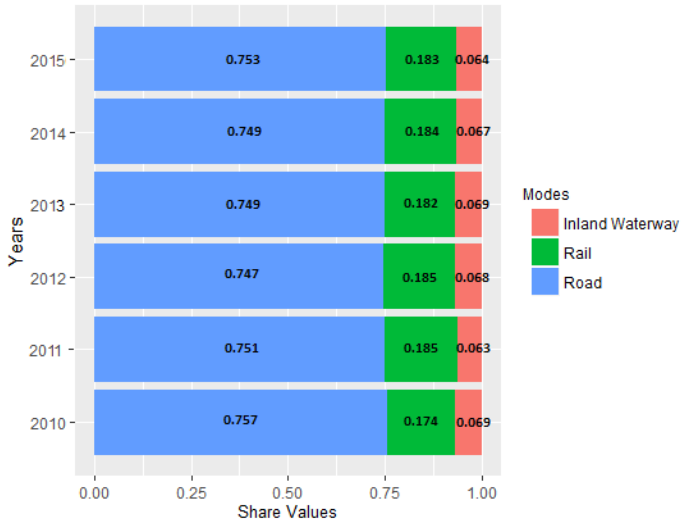
In the late seventies of the twentieth century the concept of intermodality was introduced in the transport and public policy arena (OECD, 2001). In the current literature intermodality is a widely used term accompanied by a range of different definitions. According to the objectives and the context the definition of intermodality changes. As in this research intermodality encompasses

all freight movements involving two or more modes of transportation, a broad and general definition of the OECD will be followed: *Intermodalism implies the use of at least two different modes of transport in an integrated manner in a door-to-door transport chain* (OECD, 2001).

In line with the emergence of this research and the amendment of the European Combined Transport Directive, intermodality mostly concerns transport by train, ships or barges and is served by a short road leg in the beginning and/or end of the journey. In intermodal transport the goods are loaded into intermodal loading units (usually containers) in the beginning of the journey. Loading units are moved from one type of transport to another during the journey (European Commission, 2017).

A first general overlook of the recent evolution of inland intermodal transport (by road, rail and inland waterways) follows from the EU statistics database (Eurostats, 2017). It shows that the total inland freight transport measured in tonne-kilometres increased with 1.3% during the period 2010-2015. Furthermore, the database shows that road transport continues to dominate the EU freight transport, in comparison to rail and inland waterways transport. As becomes clear from Figure 1 the share of road transport slightly decreased from 2010 to 2014, but it slightly increased from the year 2014 to 2015. In addition, the EU statistics database shows the modal split of five different transport modes: road, rail, inland waterways, air and maritime (sea). Road transport still keeps its main position in the share of transport modes, though maritime has the second greatest share. There was a rise of 1.8% in total transport of the five modes from 2010 to 2015. The total maritime transport in tonne-kilometres increased with 2.9%. The shares of maritime transport and rail transport increased with respectively 0.4% and 0.5%. Relatively, despite an absolute small rise of the road transport, the share of road transport decreased with 0.6% in share. The increased share of intermodal transport is caused by a wide bunch of factors.

Figure 1: Share of road in freight transport.



As already mentioned, there is a need for Europe to target on intermodal transport instead of transport (only) by road. The main reasons to promote this modal shift are an improvement of cost effectiveness, the support of economic growth and the reduction of social and environmental impacts (OECD, 2001). It is stated that intermodalism leads to an improvement of cost effectiveness as overall transport costs are lowered by allowing each mode to be used for that portion of the journey to which it suits best. This increasing economic productivity and efficiency enhances the nation's global competitiveness. More directly, a relative increase in intermodal transport reduces congestion, accidents, noise and the burden on over-stressed infrastructure investments. The commission states that road congestion costs 1% of the European's Gross Domestic Product (European Commission, 2016). Moreover, a relative increase in intermodal transport decreases the emission of CO2 and air pollution. Table 1 presents an oversight of the external costs for the different modalities

(Kreutzberger et al., 2003). Clearly, the costs show a significant impact of a modal shift. The European Commission states that the total negative externalities of transport create costs for society estimated at 4% of European GDP in 2011, projected to increase by around 40% by 2030 (European Commission, 2017).

Table 1: Marginal external cost per transport modality, € per 1000 tkm.

Cost Component	**Road**	**Rail**	**Inland waterway**	**Short sea**
Accidents	5.4	1.5	0	0
Noise	2.1	3.5	0	0
Air pollutions	7.9	3.8	3.0	2.0
Climate Change	0.8	0.5	Marginal	Marginal
Infrastructure	2.5	2.9	1.0	Less than 1.0
Congestion	5.5	0.2	Marginal	Marginal
Total	24.1	12.1	Maximal 5.0	Maximal 4.0
Cost difference with road traffic		*11.8*	*Ca. 19*	*Ca. 20*

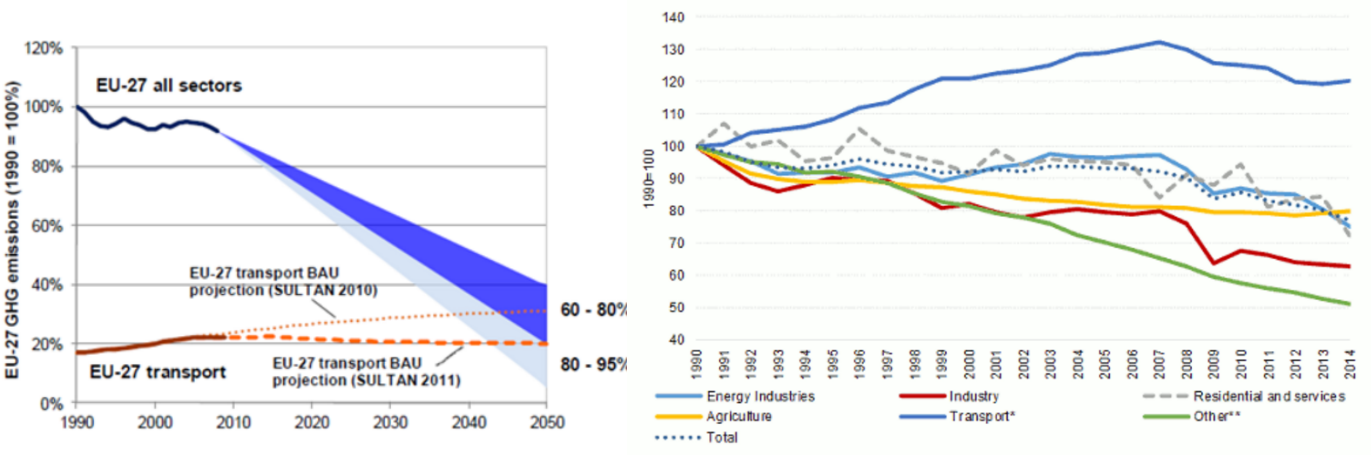
Certain policy goals are set up by European governments. In 2001, this lead to the following broadly stated objective (OECD, 2001): *The objective is to develop a framework for an optimal integration of different modes so as to enable an efficient and cost-effective use of the transport system through seamless, customer-oriented door-to-door services whilst favouring competition between transport operators.*

More specific objectives are created to construct a framework for sustainable mobility, for example in 2017 the European Commission mentioned the objective to shift 30% of road freight over 300 km to intermodal transport by 2030, and more than 50% by 2050 (European Commission, 2017). Figure 2B) shows the emission during the years of all sectors. Obviously, the transport sector had a relatively large increase in the CO2 emission. Figure 2A) shows the ambitious objective from 2010 and inwards regarding a decrease in emissions in comparison to other sectors. The objectives could result in a lot of policy actions, which for example are: tax on road transport and lower costs for rail transport (Blauwens et al., 2006), implementing an information society (European Commission, 1997) and constructing trans-European networks and nodes (TEN-T) (European Commission, 1997).

Though some of the policy actions are already implemented, the European Union still faces the complexity of the modal shift. Partly this is caused by a lower network density of intermodal transport in comparison to that of road, by a market that does not currently provide appropriate price signals to users to shift and by longer delivery times and higher costs as a result of transhipment and the complex planning of intermodal transport (European Commission, 2017). Also, in general, this is caused by different approaches of the European countries and the resulting intransparency of the intermodal transport network. Taking all these factors into account, the European Union advocates for a harmonized quantification of the intermodal transport network (OECD, 2001). A quantification will function as an important starting point to be able to fulfill the objectives. Moreover, according to the European Combined Transport Directive there is a need for new data to analyse intermodal transport. In this research we have access to a new and unique data set. We will do a quantitative analysis of the European intermodal transport flows to further unmask its grounds, evolution and causal relations. The last paragraph will continue on grounding the research in terms of concrete research questions.

First, current quantitative analyses around this topic will shortly be addressed. While existing models might be on the edge between logistical planning and econometrical analysis, there will be focused on the latter. De Jong, Gunn and Walker (2004) wrote a review of models for forecasting, policy simulation and project evaluation at the national and international levels. They distinguish four types of models that have been applied in practice: Trend and times series models, system dynamics models, zonal trip rate model and Input/Ouput models. They point out that most models are created for the sake of policy analysis on country level. Moreover, they point out that most data are aggregated on country level. Therefore, for further development they suggest models at disaggregated level. Hence, a lack of data automatically lead to the relatively short amount of

Figure 2: A) Projection of the emission objectives and B) CO2 emissions per sector.



A) B)

econometric research. Since for this research European data on the level of origin-destination could be used, a different modelling approach as well as new and deeper insights will certainly be presented.

As the need for a modal shift and the current complexities in the research field are noticed, this research tries to further clarify and therefore quantify the intermodal transport network in Europe. We will answer the following main question: "How was European intermodal freight transport organised in the past years and what factors promote a modal shift?" Additionally, more concrete research questions both content-based and methodological are formulated to further specify the direction of this research. This results into the following content-based research questions:

- How does the transport flows differ for the three modalities inland water, rail and sea?
- To what extent do socio-economic, financial and logistical factors explain the data?
- How could the differences in intermodal transport be explained geographically?

As already stated, those more concrete research questions will strengthen the main question. Furthermore, they lead to three grounds constituting the formation of data collections: A) Mode differences, B) Socio-economic, financial and logistical factors, and C) Geographical differences. Because of the novelty of both the data and econometric analyses in this research field, we find it important to critically assess its comparability and quality. That is why we raise the following methodological questions:

- To what extent do the used methods cover the research area?
- To what extent does the new and unique data set lead to new and unique results compared to the other data set?

Those questions will together strengthen the focus of this research and function as a starting point to get new insights.

3 Data

The need for further quantification of the intermodal playfield is already mentioned in the literature review. The database *Eurostat* of the European Commission and the database of the World Bank do a great job making available a lot of historic country data. For intermodal transport the availabe data from those sources is on country level. These give a good first glimpse on the evolution of the intermodal transport. However, the intermodal streams could not be scrutinized as the data lack detail[1]. In 2014 the research based consulting firm *Ecorys* started collecting data on a more disaggregated level of origin-destination, or more specifically at a origin-destination level connected to a certain terminal and carrier for transport modes rail, sea and inland water. There is data at six moments in time from 2014 to 2017. The collecting process was and still is set up step-wise and is executed manually. That's why in the latter years the data is more extensive and comprehensive.

First we will discuss some characteristics of the raw *Ecorys* data set. It contains the following variables: *Origin terminal*, *Destination terminal*, *Origin city*, *Destination city*, *Origin country*, *Destination country*, *Carrier*, *Transportation mode*, *Number of departures*, *Percentage of weekend departures* and *Transport time*. The origin-destination relation for the modes rail, sea and inland water specify an individual. For example an individual is specified as Rotterdam-Rome or Antwerpen-Berlin, by sea, rail or inland water. Mention that the individual Rotterdam-Rome differs from the individual Rome-Rotterdam (for all modes).

The variable *Number of departures* is the dependent variable. It is also mentioned as the frequency a certain origin-destination is carried out. If this variable has value 4 for individual Rotterdam-Rome by rail, it means that a train travels from Rotterdam to Rome 4 times a week.

The most important independent variables in this data set are *Transport time* and *Transportation mode*. Those cover logistical and mode-specific characteristics. Also the variable *Percentage of weekend departures* could be added in the model as independent variable. It is expected that the amount of weekend departures increases by an increasing origin-destination frequency.

The *Origin country* and *Destination country* could function as dummy variables, and later on specifically as group indicator in the specified clustered models. More importantly is that those two last mentioned variables connect the *Ecorys* data set to the data set of the World Bank (that contains country-specific data).

In Table 2 we present statistics of the dependent variable (the number of departures) of the *Ecorys* data aggregated from 2015 to 2017. Given are statistics per half year of the number of departures for all modalities and of the number of departures per modality. We aggregate data sets of different (half) years based on certain characteristics (e.g. origin-destination and carrier). According to the chosen time span the number of observations included differs. This is because different years do not contain observations with exactly the same characteristics. For example, in 2014 data from a lot of carriers are not collected, whereby a lot of origin-destinations are not included when aggregating the data sets from 2014 to 2017. Aggregation of data over the years only includes those observations that are present in all the years you want to aggregate over. Statistics of the data aggregated from 2014 to 2017 are given in Appendix A.1, together with a more extensive clarification of the whole aggregation process of the data.

Now, we briefly discuss the statistics in Table 2. It becomes clear that the mean and variance of the number of departures (or frequency) change equally through time. From 2015 inwards the mean becomes higher. This might be a result of a real increase in intermodal transport by the policy actions, climate trends or economic outcomes. Furthermore, it is clear that the percentage of observations with only one departure per week (% Frequency = 1) declined from 2015 to 2016, subsequently staying comparable. As already mentioned, for the different modalities there is a different amount of observations. Straightforwardly, the modality sea has the lowest frequency mean in comparison to the two other intermodal transport modes. This might be the result of the volume a ship over sea transports, and the time it would take, which is respectively a larger volume and a longer time.

[1] On country level we are not able to scrutinize certain factors concerning specific routes (For example, we cannot look at the relation of time and the route, we cannot select high-frequent routes only or focus on the transport corridors). Furthermore based on expert knowledge we suspect that freight data on country level is incomplete.

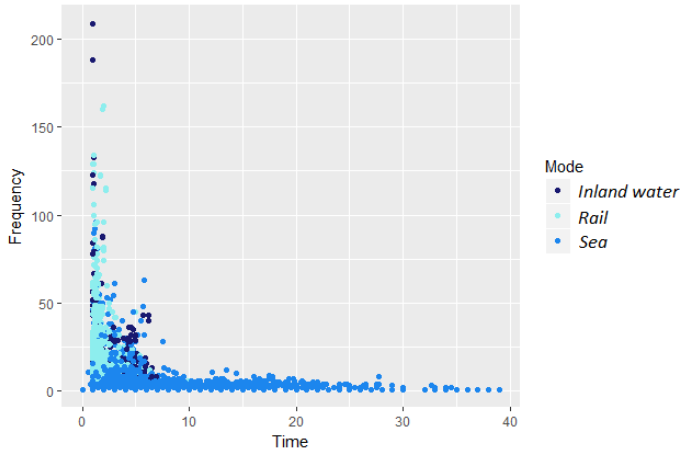
Table 2: Statistics of the number of departures per half year[a].

		$\mathbf{2015_1}$	$\mathbf{2015_2}$	$\mathbf{2016_1}$	$\mathbf{2016_2}$	$\mathbf{2017_1}$
All modalities	#obs	2401	2401	2401	2401	2401
	mean	6.43	6.51	7.35	7.09	7.11
	median	3	3	3	3	3
	mode	1	1	1	1	1
	var	134.23	140.28	190.53	171.60	172.03
	max	209	188	188	188	188
	% Frequency = 1	30.27	30.03	25.99	25.95	25.99
Rail	#obs	785	785	785	785	785
	mean	10.69	10.89	12.55	11.90	12.07
Inland water	#obs	260	260	260	260	260
	mean	13.25	13.24	14.55	15.6	15.17
Sea	#obs	1356	1356	1356	1356	1356
	mean	2.65	2.69	2.95	2.67	2.69

[a] Only those observations that are present in all half years are selected.

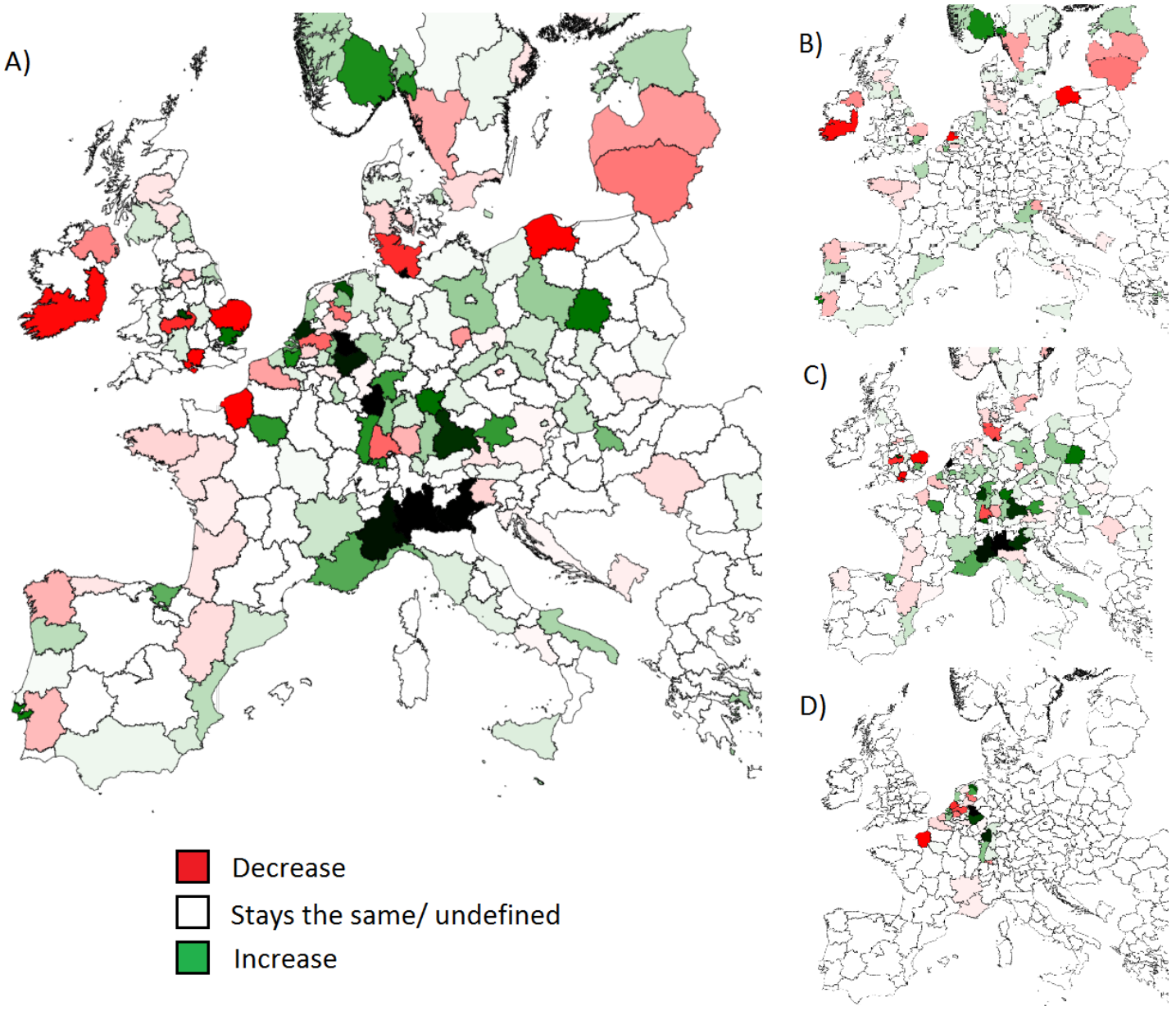
In Figure 3 this frequency-time relationship becomes clear. We see that transport that requires longer time has a lower frequency. Intuitively this sounds reasonable. Again, it is clear that transport over sea has the lowest frequency as it takes the largest time. Also, transport by inland water in comparison to transport by rail seems to be equally frequent and takes a little more time. Transport by rail is obviously fast and frequent, although of course for certain origin-destinations less frequent connections are required. This relationship will be further scrutinized in the results section.

Figure 3: The relationship of the frequency and the time certain origin-destinations comprise.



The evolution of intermodal transport based on the *Ecorys* data is visualized in Figure 4. It is clear that the intermodal transport increased during the years on the east side of the Netherlands, mainly Germany. Especially this is caused by intermodal transport by rail. Furthermore for Italy and the parts of France and Spain nearby Italy intermodal transport by rail increased (for Spain also by sea). On the other hand in the west side of France and Spain (according to the figure) the transport by train and sea diminished. Logically, transport by sea is on the shores, whereas transport by rail is inside the countries. Intermodal transport by inland water is mainly visible in the Netherlands and for a small part in Belgium and France. This might be due to the amount of canals in the Netherlands and the small distances to the sea, or either to the amount of data for intermodal transport by inland water.

Figure 4: The total increase (green) and decrease (red) of the intermodal transport frequency from 2015 to 2017 in Europe. Intermodal transport by sea, rail and inland water are respectively given by B), C) and D).



Next, we discuss the data of the World Bank. As already mentioned, the World Bank data on country level are added to the data at origin-destination level from *Ecorys*. Note that World Bank data from 2014 (2015) will be added to *Ecorys* data from 2014_2 and 2015_1 (2015_2 and 2016_1), etcetera. The following variables will be added: *GDP per capita*, *Export volume index*, *Total population*, *Total km rail lines*, *Quality of the port infrastructure* and *Pump price Diesel*. In Table 3 some statistics are presented.

Table 3: Statistics of the World Bank data from 2014-2015.

	Export volume index	GDP per capita[a]	Total population	Quality of the port infrastructure[a]	Pump price Diesel	Total km rail lines
mean	147	42694	45170500	56	169	16307
var	3205	250772347	9.621e+14	57	848	158918472
median	147	43636	60730582	57	165	15582
min	90	2114	556319	33	19	275
max	374	117507	91508084	68	250	33449

[a] There is data from 2016 available.

Next, we explain how so-called data collections will be made based on the aggregated data of *Ecorys* and the World Bank[2]. By a data collection we simply mean a specific part of the data.

Data collections are made according to different time frames and according to the three grounds (following from the three content-based research questions mentioned in the literature review): A) Mode differences, B) Socio-economic, financial and logistical factors, and C) Geographical differences. Therefore, besides different time frames, data collections are based on the inclusion of certain variables or parts of the data according to their characteristics (e.g. think of the part that only consists of data with small or large frequency). Figure 5 visualizes the factors that determine data collections. It is clear that we can assemble a huge amount of data collections, but only several will be chosen.

Different and interesting dynamics and conclusions are driven by different data collections. Furthermore, we make data collections because model-wise it is not feasible to include dummies and variables covering all relationships. Moreover, this approach enhances the oversight in the research, being able to address the results according to a structured approach. Also, not all variables are available for all years and including only a single sample (one half year) contains too less observations for all models. Lastly, note that the use of different data collections also functions as a robustness check.

Finally, we discuss the data of *Eurostat*[3]. It concerns data aggregated on country level with an earlier and larger time span (1995-2013). The dependent variables from *Eurostat* are *Inland water freight quantity*, *Rail freight quantity* and *Road freight quantity*. The following independent variables are included: *GDP per capita indexed*, *Population*, *Transport investments*, *Trade index*, *Environment tax index*, and *Policy climate stringency index*. Statistics are presented in Table 4.

It comes down to a panel data set of 20 countries[4] and 19 years. There is a missing data point in the year 2007 for the rail freight quantity variable. This will be imputed by taking the mean over the years 2006 and 2008.

In the Appendix A.9 we show the change of the amount of rail, road and intermodal freight transported in tonne-km from 2000 to 2013. We selected these time points since from 2000 there is active policy aimed at an increase of intermodal transport. As is visible, the amount of intermodal transport did not increase for all countries.

In the Appendix A.10 we show the shares of the countries for the above mentioned variables, given are the shares for 1995 and 2013.

[2]Data collections will be used in fixed and random effects models.

[3]*Eurostat* data will be used in the market share models.

[4]We will analyze the following countries: Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Luxembourg, Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland and United Kingdom.

Figure 5: Several data collections could be made according to characteristics to be analyzed.

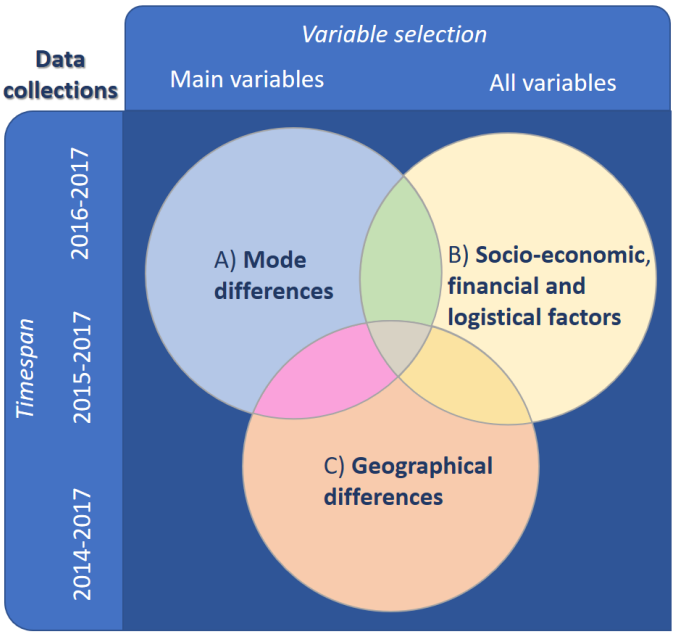


Table 4: Statistics of *Eurostat* data from 1995-2013[a,b].

	Road freight	Rail Freight	Inl. water freight[c]	GDP per cap.	Population	Transport investments	Trade index	Environ. tax index	Climate index
mean	87795	20720	7571	102	27781536	5.70e+9	234	2.50	2.09
var	7.33e+09	5.57e+08	3.07e+08	1599.93	6.66e+14	3.70e+19	69370	0.27	0.68
min	2563	212	25.10	37.31	408625	132996307	6.95	1.38	0.68
max	335868	117382	66465	248	82534180	2.33e+10	1461	3.81	4.13

[a] This data will get log-transformed later.
[b] Freights are given in million tonne-km and GDP per Capita is given in index-form.
[c] Note that for Inland water, there is no data from the Netherlands, Italy, Spain and Poland.

4 Methods

In this section we explain and discuss the used methods. First we explain fixed and random effects models that are applied on the (*Ecorys* and World Bank) data on origin-destination level. We want to declare the number of departures by (possibly) related variables. Both linear and count models will be addressed. Also the clustered version of those models will be discussed. Then we explain the market share model that will be applied on *Eurostat* data on country level as a case study. Finally, we briefly address the statistics.

4.1 Fixed and random effects models

As already mentioned, in this research we use a panel data set from *Ecorys* at a level of origin-destination (aggregated is data from the World Bank). The following independent variables are used: *Transport time*, *Transportation mode*, *Percentage of weekend departures*, *GDP per capita*, *Export volume index*, *Total population*, *Total km rail lines*, *Quality of the port infrastructure* and *Pump price Diesel*. An advantage of panel data is the possibility of mapping dynamics of individual behaviour (Cameron, 2005). For example, it can be determined if a decrease in length of a journey (through the time) leads to an increase of the frequency.

Furthermore, a major advantage of panel data is the possibility of allowing for unobserved individual heterogeneity, in other words it allows each cross-sectional unit to have a different intercept. This leads to a individual-specific effects model. Treating unobserved individual heterogeneity as correlated with regressors or as being distributed independently of the regressors results respectively in fixed effects and random effects models. We find this interesting, as the latent part (unobserved individual heterogeneity) in modelling the frequency of a certain origin-destination might be either fixed or random in specific data sets. The quantification of the latent part is valuabe for comparison reasons. Is the trip from Rotterdam to Rome more frequently executed than the the trip from Rotterdam to Barcelona? And could the variables explain this difference or is there a latent part, random or fixed?

Additionally, the fixed and random effects also can be addressed at group level. Fixed and random effect comparisons could be made on country or TEN-T level (promoted routes by the European Union).

Altogether, this leads to linear (grouped), Poisson (grouped) and Negative Binomial fixed and random effects models.

4.1.1 Linear models

For comparison we will first apply an OLS regression. Clearly, this results in an overall constant that does not cover individual unobserved heterogeneity. Although this is the most parsimonious model, as the amount of parameters to be estimated is small, the individual-specific models are more interesting, because we expect that a large part of the number of departures is explained by unobserved heterogeneity.

The individual-specific effects model is formulated as follows (Cameron and Trivedi, 2005):

$$y_{it} = \alpha_i + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \epsilon_{it}, \tag{1}$$

The α_i's capture unobserved heterogeneity. As mentioned before, in case of a fixed effects model α_i will be treated as an unobserved variable that is potentially correlated with the observed regressors. In this case the α_i's need to be estimated. The alternative model is the random effects model. Here the α_i's are distributed independently of the regressors:

$$\begin{aligned} \alpha_i &\sim (\alpha, \sigma^2_\alpha) \\ \epsilon_{it} &\sim (0, \sigma^2_\epsilon). \end{aligned} \tag{2}$$

Parameters are estimated differently for both fixed and random effects models.

First we take a look at the fixed effects model. We will apply the so-called Within method (leading to $\boldsymbol{\beta}_W$), because it is the most efficient for $T > 2$. We follow the execution from the book of Cameron and Trivedi (1998, page 726).

The distribution of the estimator for $\boldsymbol{\beta}_W$ are required for deriving the significance of the causal relations in question. The usual OLS results do not apply, because the error terms are correlated over time. Hence a robust sandwich estimator has to be used, see Cameron and Trivedi (1998, page 727).

Next, we discuss the random effects model. In the data set some variables are time-invariant. Fixed effects models are not able to estimate those parameters. As random effects models can estimate parameters for time-invariant variables, these are valuable for the analyses of the causal relationships. Besides this advantage, the unobserved heterogeneity might be random in certain data collections. Furthermore, the differences between fixed and random parameter estimates are interesting.

The random effects estimator of μ (the intercept) and β is also called the GLS estimator, which we will estimate by OLS regression of the following transformed model:

$$y_{it} - \hat{\lambda}\bar{y_i} = (1-\hat{\lambda})\mu + (\boldsymbol{x}_{it} - \hat{\lambda}\bar{\boldsymbol{x}}_i)'\boldsymbol{\beta}_{RE} + (1-\hat{\lambda})\alpha_i + (\epsilon_{it} - \hat{\lambda}\bar{\epsilon_i}), \tag{3}$$

where $\bar{y_i}$, $\bar{x_i}$ and $\bar{\epsilon_i}$ are respectively the means of $(y_{i1} \dots y_{iT})$, $(x_{i1} \dots y_{1T})$ and $(\epsilon_{i1} \dots \epsilon_{1T})$. $\hat{\lambda}$ is a consistent estimator for $\lambda = 1 - \frac{\sigma_\epsilon}{(T\sigma^2_\alpha + \sigma^2_\epsilon)^{1/2}}$. In comparison to the fixed effects models, an intercept

is introduced so that the random effects can be normalized to have zero mean. In Appendix A.3 we elaborate on how to estimate the unknown components of λ.

For short panels the robust estimate of the asymptotic variance yields (for general formulation of this estimate see Appendix A.2):

$$V[\hat{\boldsymbol{\beta}}_{RE}] = \Big[\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{\boldsymbol{x}}_{it}\tilde{\boldsymbol{x}}'_{it}\Big]^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{\boldsymbol{x}}_{it}\tilde{\boldsymbol{x}}'_{it}\hat{\tilde{\epsilon}}_{it}\hat{\tilde{\epsilon}}_{it}\Big[\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{\boldsymbol{x}}_{it}\tilde{\boldsymbol{x}}'_{it}\Big]^{-1}, \tag{4}$$

where $\tilde{\boldsymbol{x}}_{it} = (\boldsymbol{x}_{it} - \hat{\lambda}\bar{\boldsymbol{x}}_i)$ and $\tilde{\epsilon}_{it} = (\epsilon_{it} - \hat{\lambda}\bar{\epsilon_i})$.

Now, we have clarified the linear individual-specific fixed and random effects models.

We find it interesting to analyze the unobserved heterogeneity of certain European countries or TEN-T corridors. Therefore, we introduce the clustered fixed and random effects models (Cameron and Trivedi, 2005):

$$y_{itc} = \alpha_c + \boldsymbol{x}'_{itc}\boldsymbol{\beta} + \epsilon_{itc}, \tag{5}$$

here i= 1...N_c, c= 1...C, and t= 1...T. Cameron and Trivedi (2005) only distinguish the panel data case in which units are observed more than once (notation: it) and the cluster-specific case where units are observed only once (notation: ic). In our case the data units are both observed through time and in clusters. Since for both the panel data case and the cluster-specific case the setup and terminology are parallel, we use an analogous approach for the combination of both. For feasibility sake, we consider the observations over time as being separate individuals, but in the same cluster. There is no effect on the estimation of the $\boldsymbol{\beta}$ parameters as well as on the parameters of unobserved heterogeneity.

We will use the Within method for the (panel) cluster-specific fixed effects case either. The $\boldsymbol{\beta}$ estimates follow from the Within method as given by Cameron and Trivedi (1998, page 840). The cluster-robust (co)variance matrix will be calculated in a parallel manner as in the individual-specific fixed effects case.

For the clustered random effect model we again use the GLS method. Comparably to the individual-specific case this leads to estimation by OLS regression of the transformed model from (3), see Cameron and Trivedi (1998, page 837). The cluster-robust (co)variance matrix will be calculated in a parallel manner as in the individual-specific random effects case.

Now we have discussed all linear models, we will briefly mention the coefficient interpretation. The α_i's cover the unobserved heterogeneity. In the fixed effects case they also cover the overall intercept. The other parameters ($\boldsymbol{\beta}$) are interpreted as follows: by a one unit increase in an independent variable, the dependent variable increases by the corresponding β.

4.1.2 Poisson models

Since the real data is count data, the linear models might not perform optimally. Therefore, we introduce Poisson fixed and random effects models and follow the set-up of both Cameron and Trivedi (1998) and Hausman et al. (1984).

The following is assumed about the distribution of the number of departures y_{it}:

$$\begin{aligned} y_{it} &\sim \mathrm{P}[\mu_{it} = \alpha_i\lambda_{it}], \\ \lambda_{it} &= \exp(\boldsymbol{x}'_{it}\boldsymbol{\beta}). \end{aligned} \tag{6}$$

Note that λ_{it} and μ_{it} are used in a different way as in the previous chapter and that α_i still refers to the individual effect. Then, for a single observation y_{it} the density is:

$$f(y_{it}|\boldsymbol{x_i}, \alpha_i, \boldsymbol{\beta}) = \exp(-\mu_{it})(\mu_{it})^{y_{it}}/y_{it}!. \tag{7}$$

Furthermore, from the standard Poisson characteristics we know that the conditional mean is:

$$\begin{aligned} \mathrm{E}[y_{it}|\boldsymbol{x}_{it},\alpha_i] &= \mu_{it} \\ &= \alpha_i \exp(\boldsymbol{x}'_{it}\boldsymbol{\beta}) \\ &= \exp(\delta_i + \boldsymbol{x}'_{it}\boldsymbol{\beta}). \end{aligned} \tag{8}$$

Note here that both an additive and a multiplicative individual-specific effects model are presented. Also, the conditional mean equals the conditional variance: $\mathrm{E}[y_{it}|\boldsymbol{x}_{it},\alpha_i] = \mu_{it} = \mathrm{Var}[y_{it}|\boldsymbol{x}_{it},\alpha_i]$. In the data section we noted that (for most data collections) the variance of the number of departures is larger than its mean.This leaves space for further consideration. We follow the conventional approach to (so-called) overdispersion (Allison and Waterman, 2002) and use robust standard errors.

First, we present the Poisson fixed effects model. The Poisson maximum likelihood function simultaneously estimates $\boldsymbol{\beta}$ and all α_i's. Based on equation (7) the log-likelihood is:

$$\begin{aligned} \ln \mathrm{L}(\boldsymbol{\beta},\alpha_i) &= \ln\Big[\prod_i \prod_t \exp(-\alpha_i\lambda_{it})(\alpha_i\lambda_{it})^{y_{it}}/y_{it}!\Big] \\ &= \sum_i \Big[-\alpha_i \sum_t \lambda_{it} + \ln\alpha_i \sum_t y_{it} + \sum_t y_{it}\ln\lambda_{it} - \sum_t \ln y_{it}!\Big]. \end{aligned} \tag{9}$$

If this log-likelihood is differentiated with respect to α_i and logically set to zero for its maximum, we get: $\hat{\alpha_i} = \sum_t y_{it} / \sum_t \lambda_{it}$. A substitution of $\hat{\alpha_i}$ into (9) drops the α_i's and leads to the following concentrated likelihood function:

$$\begin{aligned} \ln \mathrm{L}_{concentrated}(\boldsymbol{\beta}) &= \sum_i \Big[-\sum_t y_{it} + \ln\frac{\sum_t y_{it}}{\sum_t \lambda_{it}} \sum_t y_{it} + \sum_t y_{it}\ln\lambda_{it} - \sum_t \ln y_{it}!\Big] \\ &\propto \sum_i \sum_t \Big[y_{it}\ln\lambda_{it} - y_{it}\ln\big(\sum_s \lambda_{is}\big)\Big]. \end{aligned} \tag{10}$$

Cameron and Trivedi (2005) state that the estimation of this Poisson fixed effects model provide consistent estimates of $\boldsymbol{\beta}$. The following estimator for panel-robust (co)variance is used (for general formulation of this estimate see Appendix A.2):

$$V[\hat{\boldsymbol{\beta}}_{PoisFE}] = \Big[\sum_{i=1}^{N}\sum_{t=1}^{T} \hat{\lambda}_{it}\boldsymbol{x}_{it}\boldsymbol{x}'_{it}\Big]^{-1} \sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} \boldsymbol{x}_{it}\boldsymbol{x}'_{it}\hat{\epsilon}_{it}\hat{\epsilon}_{it} \Big[\sum_{i=1}^{N}\sum_{t=1}^{T} \hat{\lambda}_{it}\boldsymbol{x}_{it}\boldsymbol{x}'_{it}\Big]^{-1}, \tag{11}$$

where $\hat{\epsilon}_{it} = y_{it} - \hat{\lambda}_{it}$.

Now, as we did for the linear case either, we discuss a random effects model. In general, for a Poisson random effects model the random effects are assumed to be gamma-distributed. In Cameron and Trivedi (2015) α_i is distributed $G[\theta_1,\theta_1]$, which automatically leads to mean is 1, variance is $\frac{1}{\theta_1}$, and density $\mathrm{g}(\alpha_i|\theta_1) = \theta_1^{\theta_1}\alpha_i^{\theta_1-1}\exp(-\alpha_i\theta_1)/\Gamma(\theta_1)$. The distribution parameters are limited in the sense that the rate and the shape parameter are equal. In general this would not be seen as a limitation to the outcome, because the estimations of $\boldsymbol{\beta}$ are more important. As this research also questions the size of the latent variable, in other words the unobserved heterogeneity, α_i is set to be distributed $G[\theta_1,\theta_2]$, which automatically leads to a mean of $\frac{\theta_1}{\theta_2}$, a variance of $\frac{\theta_1}{\theta_2^2}$ and distribution $\mathrm{g}(\alpha_i|\theta_1,\theta_2) = \theta_2^{\theta_1}\alpha_i^{\theta_1-1}\exp(-\alpha_i\theta_2)/\Gamma(\theta_2)$.

The Poisson model (7) with gamma-distributed random effects density $\mathrm{g}(\alpha_i|\theta_1,\theta_2)$ results to the unconditional joint density for observation i:

$$f(\boldsymbol{y}_i|\boldsymbol{X}_i,\boldsymbol{\beta},\boldsymbol{\theta}) = \int \prod_t \Big[\exp(-\alpha_i\lambda_{it})(\alpha_i\lambda_{it})^{y_{it}}/y_{it}!\Big]\theta_2^{\theta_1}\alpha_i^{\theta_1-1}exp(-\alpha_i\theta_2)/\Gamma(\theta_2)\delta\alpha_i. \tag{12}$$

Since an analytic solution for this integral is infeasible, the integral will be approximated by the numerical Gauss–Hermite quadrature method. It approximates the integral with respect to a normal density by a weighted sum. A general set-up and some further explanation regarding this method is given in Appendix A.4. For this specific case the result of the log likelihood is given in Appendix A.5. Following Cameron and Trivedi (2005) we can use the panel-robust standard errors of (11).

Like mentioned for linear models, in this research we find it interesting to analyze unobserved heterogeneity of certain clusters. Similar to the linear case, for feasibility sake, we take the observations over time to being separate individuals, but in the same cluster. For Poisson clustered fixed effect models, based on (7) and comparable to (9), we get the following log likelihood:

$$\ln \mathrm{L}(\boldsymbol{\beta}, \alpha_i) = \cdots = \sum_c \Big[-\alpha_c \sum_i \lambda_{ic} + \ln\alpha_c \sum_i y_{ic} + \sum_i y_{ic} \ln\lambda_{ct} - \ln y_{ic}! \Big]. \tag{13}$$

Following Cameron and Trivedi (2005), unlike the linear model, here it is not possible to eliminate the incidental parameters $\alpha_1, \ldots, \alpha_C$. Therefore we estimate the incidental parameters simultaneously to the β parameters. Since the cluster-based approach is much more parsimonious in its incidental parameters this is usually more feasible. Note that the cluster-robust (co)variance matrix will be calculated in a parallel manner as in the Poisson individual-specific fixed effects case.

As Cameron and Trivedi (2005) state, for a clustered random effects model, the estimates of the random effects Poisson case (resulting from (40) in Appendix A.5) are consistent. Besides the potential for small efficiency gains, there is no reason to adjust those estimators for the clustered case. Therefore we use the estimates of the random effects Poisson model.

4.1.3 Negative binomial models

As already mentioned, the individual effects Poisson model is restricted because the conditional variance should equal the conditional mean. In the data section it became clear that (for most data collections) the variance of the dependent variable is larger than the mean. The model does not account for some additional heterogeneity; which Allison and Waterman (2002) refer to as the problem of overdispersion. The problem of overdispersion might be resolved by the use of robust standard errors. However, this does not solve the problem of the model mis-specification. Therefore we introduce another fixed and random effects count model: The negative binomial model.

First we discuss the negative binomial fixed effects model, conform the terminology of the Poisson model. There are several different approaches to the negative binomial fixed effects model[5]. We discuss the so-called NB2 variant of Allison and Waterman (2002).

The NB2 negative binomial model is based on the following mass function:

$$f(y_{it}|\mu_{it}; o_i) = \frac{\Gamma(o_i + y_{it})}{\Gamma(o_i)\Gamma(y_{it}+1)} \Big(\frac{\mu_{it}}{\mu_{it}+o_i}\Big)^{y_{it}} \Big(\frac{o_i}{o_i+\mu_i t}\Big)^{o_i}. \tag{14}$$

Here o_i is the so-called overdispersion parameter. Furthermore the mean is μ_{it} and the variance is $\mu_{it}(1 + \mu_{it}/\lambda_{it})$. Following the assumption that event counts are independent across time for each individual, there is no manageable conditional likelihood. This is because there is no complete sufficient statistic for the α_i's (or in additive terminology as stated in equation (8), no statistic for the δ_i's). To avoid this problem we simply estimate α_i together with the other parameters, including o_i. In this research o_i is assumed to be constant for all individuals, which leads to the overall overdispersion parameter o (otherwise there would be too many parameters). Continuing on (14), and restricting o_i to be the same for all individuals, we maximize the following log likelihood:

$$\sum_i \sum_t \Big[\log\Big(\frac{\Gamma(o+y_{it})}{\Gamma(o)\Gamma(y_{it}+1)}\Big) + y_{it}\log\Big(\frac{\mu_{it}}{\mu_{it}+o}\Big) + o\,\log\Big(\frac{o}{o+\mu_{it}}\Big)\Big]. \tag{15}$$

Unfortunately, the practical problem of computational infeasability still appears, caused by too many dummy variables (equaling to the amount of individuals). We circumvent this problem by using the α_i estimators from the Poisson distribution as starting values.

[5] In 1984 Hausman, Hall and Griliches introduced a certain negative binomial model (a so-called NB1 model). Allison and Waterman (2002) notice that their formulation of the negative binomial model is not a true fixed effects model, because the α_i's play a different role than x_{it}. Applying this model to our data did not result in reasonable outcomes.

Allison and Waterman (2002) also present the multivariate generalization of the negative binomial distribution as an alternative. This formulation does not concern overdispersion and implementation leads to exactly the same outcome as the Poisson estimators. It is of no addition in our quest for an appropriate negative binomial model.

Previously, for both Linear and Poisson models, the (co)variance needed to be robust. For the negative binomial fixed effects model this is not necessary, as it already takes into account overdispersion. Therefore, we use (32) or (33) from Appendix A.2. A derivation of both are presented in Appendix A.6.

Next, we will discuss the random effects negative binomial model. Fortunately there are less approaches than for the fixed effects case. We follow the approach of Hausman, Hall and Griliches (1984). The incidental parameters α_i are beta distributed. Hausman, Hall and Griliches (1984) integrate the assembly of the negative binomial distribution and the beta distribution which results in the following individual specific mass distribution:

$$f(y_{i1} \dots y_{it} | x_{i1} \dots x_{it}) = \frac{\Gamma(a+b)\Gamma(a+\sum_t \lambda_{it})\Gamma(b+\sum_t y_{it}))}{\Gamma(a)\Gamma(b)\Gamma(a+b+\sum_t \lambda_{it}+\sum_t y_{it})} \prod_t \frac{\Gamma(\lambda_{it}+y_{it})}{\Gamma(\lambda_{it})\Gamma(y_{it}+1)}. \tag{16}$$

This leads to the following total log likelihood function to maximize:

$$\begin{aligned} \sum_i \Big[& \log\Gamma(a+b) + \log\Gamma(a+\sum_t \lambda_{it}) + \log\Gamma(b+\sum_t y_{it})) \\ & - \Big(\log\Gamma(a) + \log\Gamma(b) + \log\Gamma(a+b+\sum_t \lambda_{it} + \sum_t y_{it}) \Big) + \\ & \sum_t \Big(\log\Gamma(\lambda_{it}+y_{it}) - \log\Gamma(\lambda_{it}) - \log\Gamma(y_{it}+1) \Big) \Big]. \end{aligned} \tag{17}$$

In the paper of Hausman, Hall and Griliches (1984) the (co)variance estimator is not given. Therefore we take the other models as benchmarks for the significance of the parameters.

In comparison to the linear models, the coefficient interpretation of the Poisson and negative binomial model is more complicated. In both models the mean of the dependent variable $E[y_{it}|\boldsymbol{x}_{it}]$ comes down to $\alpha_i \exp(\boldsymbol{x}'_{it}\boldsymbol{\beta})$ (or in additive notation $\exp(\delta_i + \boldsymbol{x}'_{it}\boldsymbol{\beta})$). The α_i's cover the unobserved heterogeneity (and the overall intercept). A one unit increase in this part would lead to an increase of $\exp(\boldsymbol{x}'_{it}\boldsymbol{\beta})$ in the dependent variable. The effect of the unobserved heterogeneity thus depends on the value of the other independent variables and their corresponding parameter value.

Even more interesting are the marginal effects of a one-unit increase of a certain variable, or in other words: the changes in $E[y_{it}|\boldsymbol{x}_{it}]$ due to a change in variable $x_{it,j}$. This marginal effect equals:

$$\frac{\delta \mathrm{E}[y_{it}|\boldsymbol{x}_{it}]}{\delta x_{it,j}} = \exp(\boldsymbol{x}'_{it}\boldsymbol{\beta})\beta_j. \tag{18}$$

β_j determines the sign of the marginal effects, as $\exp(\boldsymbol{x}'_{it}\boldsymbol{\beta})$ is always positive.

A more usual interpretation is the elasticity, which is a percentage change in $E[y_{it}|\boldsymbol{x}_{it}]$ due to a percentage change in $x_{it,j}$. This is given by:

$$\frac{\delta \mathrm{E}[y_{it}|\boldsymbol{x}_{it}]}{\delta x_{it,j}} \times \frac{x_{it,j}}{\mathrm{E}[y_{it}|\boldsymbol{x}_{it}]} = \beta_j x_{it,j}. \tag{19}$$

4.2 Market share model

In this section we introduce the market share attraction model. This model is valuable for describing (and forecasting) market shares. We apply it to the country specific data from *Eurostat* to compare the positions of several countries in the freight market[6]. We use freight country shares for road, rail and inland water transport. For the sake of ease and clarity, we only estimate the own effects of the variables (*GDP per capita indexed*, *Population*, *Transport investments*, *Trade index*, *Environment tax index*, and *Policy climate stringency index*). For the econometric analysis we follow the paper of Fok, Franses and Paap (2002). Note that the terminology of the fixed and random effects models is different.

[6]We use the market share model since (in addition to the fixed and random effects models) we are more interested in describing the freight amount relative to other countries than the absolute freight amount.

We will first discuss a representation of a general market share model. Then the estimation methods will be presented. Finally, we shortly clarify the parameter interpretation.

We implement a restricted version of the general market share model. This restriction will be both on competition and dynamics. Therefore, let A_{it} be the so-called attraction (in our case the amount of freight transport) from country i at time t, $t = 1,\ldots,$T, given by:

$$A_{it} = \exp(\mu_i + \epsilon_{it}) \prod_{k=1}^{K} x_{k,it}^{\beta_{k,i}} \text{ for } i = 1, \ldots, I, \tag{20}$$

where $x_{k,i,t}$ stands for the k-th explanatory variable (such as GDP, export and import, CO2 emissions, etc.) for country i at time t and where $\beta_{k,i}$ is the corresponding coefficient, which in this case is assumed to be constant for all countries (this leads to β_k). Further, the parameter μ_i is country-specific and the error terms are normally distributed with zero mean and (co)variance matrix Σ.

We follow the market share theorem, which says that the market share of country i at time t is equal to its attraction divided by the sum of all attractions:

$$M_{it} = \frac{A_{it}}{\sum_{j=1}^{I} A_{jt}}. \tag{21}$$

The model in (20) is linearized to enable parameter estimation. The following two steps are followed. First, we choose country I as a benchmark country (United Kingdom). This leads to:

$$\begin{aligned} \frac{M_{it}}{M_{It}} &= \frac{A_{it}}{\sum_{j=1}^{I} A_{jt}} \Big/ \frac{A_{It}}{\sum_{j=1}^{I} A_{jt}} = \frac{A_{it}}{A_{It}} \\ &= \frac{\exp(\mu_i + \epsilon_{it}) \prod_{k=1}^{K} x_{k,it}^{\beta_{k,i}}}{\exp(\mu_I + \epsilon_{It}) \prod_{k=1}^{K} x_{k,It}^{\beta_{k,I}}}. \end{aligned} \tag{22}$$

Next, in the second step we take the natural logarithm. This leads to the (I-1)-dimensional set of equations, given by:

$$\log M_{it} - \log M_{It} = (\mu_i - \mu_I) + (\epsilon_i - \epsilon_I) + \sum_{k=1}^{K} \Big(\beta_{k,i} \log\, x_{k,it} - \beta_{k,I} \log\, x_{k,It} \Big). \tag{23}$$

Now, we discuss the parameter estimation. Although Fok, Franses and Paap (2002) also discuss a different estimation method, we implement the one with a base brand. Two approaches will be used: OLS and GLS.

The set of equations from (23) is rewritten as follows:

$$\begin{aligned} y_{1t} &= w_{1t}' \boldsymbol{b}_1 &&+ z_{1t}' \boldsymbol{a} &&+ \eta_{1t} \\ \vdots &= \vdots &&+ \vdots &&+ \vdots \\ y_{(I-1)t} &= w_{(I-1)t}' \boldsymbol{b}_{I-1} &&+ z_{(I-1)t}' \boldsymbol{a} &&+ \eta_{(I-1)t}. \end{aligned} \tag{24}$$

Here $y_{it} = \log M_{it} - \log M_{It}$, $\boldsymbol{\eta_t}' \sim \text{NID}(\mathbf{0}, \hat{\Sigma})$. w_{it} and z_{it} are k-dimensional vectors of explanatory variables with regression coefficient vector $\boldsymbol{a}$. For parameter estimation a matrix notation is preferred: $y_i = (y_{i1}, \ldots, y_{iT})'$, $\boldsymbol{Z}_i = (z_{i1}, \ldots, z_{iT})'$ and $\eta_i = (\eta_{i1}, \ldots, \eta_{iT})'$ for $i = 1, \ldots, (I-1)$. We get:

$$\begin{pmatrix} y_{1t} \\ \vdots \\ y_{(I-1)t} \end{pmatrix} = \begin{pmatrix} \boldsymbol{W}_1 & \mathbf{0} & \ldots & \mathbf{0} & \boldsymbol{Z}_1 \\ \mathbf{0} & \boldsymbol{W}_1 & \ldots & \mathbf{0} & \boldsymbol{Z}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \boldsymbol{W}_{I-1} & \boldsymbol{Z}_{I-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{b}_{I-1} \\ \boldsymbol{a} \end{pmatrix} + \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_{(I-1)} \end{pmatrix}, \tag{25}$$

which can be stated as:

$$y = Xb + \eta, \tag{26}$$

where $\eta \sim N(\mathbf{0}, (\hat{\Sigma} \otimes \boldsymbol{I}_T))$. As Fok, Franses and Paap (2002) point out, OLS estimates are consistent and either efficient if the explanatory variables are equal for each equation. As this might not be the case in this research, OLS might give inefficient estimates.

Also we estimate the feasible GLS estimator (by the SUR method). The iterative SUR estimator uses the OLS estimates as its starting point to estimate the (co)variance matrix ($\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} \hat{\eta}_t \hat{\eta}_t'$) and is constructed as follows:

$$\hat{b}_{SUR} = (X'(\hat{\Sigma}^{-1} \otimes \boldsymbol{I}_T)X)^{-1} X'(\hat{\Sigma}^{-1} \otimes \boldsymbol{I}_T)y, \tag{27}$$

where $\hat{\eta}_t$ (for $\hat{\Sigma}$) results from stacking $\hat{\eta}_{it} = y_{it} - w_{it}'\hat{\boldsymbol{b}}_{SUR} - z_{it}'\hat{\boldsymbol{a}}_{SUR}$. Those will be used to estimate the (co)variance matrix again. This procedure will be iterated until convergence of both estimates.

The standard errors for the estimates regression parameters b are estimated as follows:

$$\hat{V}(\hat{b}) = (X'(\hat{\Sigma}^{-1} \otimes \boldsymbol{I}_T)X)^{-1}. \tag{28}$$

Note that a sufficiently large T to estimate the covariance of the $\hat{\eta}_t$ is needed. It might lead to a wrong outcome for the significance of the parameters.

Comparable to the Poisson and negative binomial fixed and random effects models, a coefficient interpretation based on the elasticity is easiest. For model (20) the elasticity of the k-th coefficient is (for a derivation of this result, see Appendix A.7):

$$\frac{\delta M_{it}}{\delta x_{k,jt}} \frac{x_{k,jt}}{M_{it}} = (\delta_{i=j} - M_{jt})\beta_k. \tag{29}$$

where $\delta_{i=j}$ is 1 if i equals j and 0 if not. This rationale presents the percentage change in M_{it} due to a percentage change in $x_{k,jt}$.

4.3 Statistics

For the data set on origin-destination level we explained different fixed and random effects models in detail. Since all models differ in their theoretical advantages and disadvantages, it is hard to conclude from the theory which model fits this data set best, let alone comparisons of the best fits for the different data collections. Therefore, we compare the models based on a test and performance measures. The Hausman test tests if the model is random or fixed. As performance measures the mean squared error (MSE), mean absolute deviation (MAD), mean absolute percentage error (MAPE) and the likelihood will be used. Lastly, the variable selection procedure will shortly be mentioned.

The Hausman test more specifically tests if fixed effects are present (Cameron and Trivedi, 2005). It tests whether the fixed and random effects estimators are statistically significantly different. It assumes that the random effects model is the true model. This leads to the null hypothesis which says that the individual-specific effects are uncorrelated with regressors. If rejected, we can conclude that fixed effects are present.

Following Cameron and Trivedi (2005), we assume a random effects model with α_i is iid $(0, \sigma_\alpha^2)$ uncorrelated with the regressors and errors ϵ_{it} iid $(0, \sigma_\epsilon^2)$. It is assumed that the random effects estimators are fully efficient. The Hausman test is given by:

$$\mathrm{H} = \left(\hat{\beta}_{T,RE} - \hat{\beta}_{FE}\right)' \left[\hat{V}[\hat{\beta}_{FE}] - \hat{V}[\hat{\beta}_{T,RE}]\right]^{-1} \left(\hat{\beta}_{T,RE} - \hat{\beta}_{FE}\right), \tag{30}$$

where $\hat{\beta}_{T,RE}$ refers to the time-varying estimated parameters. This test statistic has a $\chi^2(dim(\beta_{FE}))$-distribution under the null hypothesis.

In most cases the errors will not be iid and so the RE-estimator is not fully efficient. That is why we implemented a bootstrap approach, but this gave unreasonable results. Therefore, we only

present results for the Hausman test presented above. It is generally known that the definition of the Hausmann test statistics results in negative values, which is in conflict with the χ^2- statistics. Schreiber (2008) states that taking the absolute value of the test statistic is the remedy to this problem as it leads to comparable conclusions.

Next, we discuss model performance. Model prediction errors reflect the model performance. Model prediction is quite straightforward; We follow the model specification in the method section and include the regressors in question. For the random effects model the α_i for prediction is calculated by $E[\alpha_i|y_{i1} \dots y_{iT}, \mathbf{x}_i; \theta]$. We explain the calculation of this expectation in Appendix A.8. Many different loss functions could be suggested to estimate the prediction errors. As the data seem to be without outliers, the MSD, MAPE and MAD loss functions are appropriate and easy to interpret. Note that those loss functions only inform about the model parameters and their performance; They do not inform about the stochastic characteristics of the data, or in other words, the distribution fit.

The MSD is a very common loss function. It simply measures the difference between the real value and the estimation of that value, as follows:

$$\text{MSE} = \frac{1}{NT} \sum_{i}^{N} \sum_{t}^{T} (y_{it} - \hat{y}_{it})^2, \tag{31}$$

where $\hat{y}_{it}$ logically depends on the model in question. Note that large differences will be emphasized by the quadratic. A comparable loss function is the MAD. It expresses the loss in units of the data whereby it explodes large errors less than the MSD does. It is calculated by replacing the quadratic of the difference in (31) by the absolute value of the difference. Another comparable loss function is the MAPE. The interpretation is different as the prediction accuracy is measured in percentages. It is calculated by replacing the quadratic of the difference in (31) by the absolute value of this difference divided by y_{it}.

A final performance measure is the (log) likehood, which is a contribution to the just mentioned loss functions, as it informs about the distribution fit.

For the fixed and random effects models we select independent variables in different data collections based on the p-values of all models and on the condition that the models are executable. Mention that for different models different variables are significant and that selection is based on all models. For the market share models we select the independent variables based on the p-values (for all different dependent variables).

5 Results

In this section we address the results of the research. We will present the results from the point-of-view of the main question and the content-based research questions. Thus, keeping in mind Figure 5, the results will be presented according to the following three grounds: A) Mode differences, B) Socio-economic, financial and logistical factors, and C) Geographical differences. We present results of a general model based only on *Ecorys* data. Then, we will discuss results for the models addressing the three just-mentioned grounds.

The different models and data collections obviously lead to different results. Note that different models involve different time frames, because not all variables are available for all years. Furthermore, using one sample for all models involves too less observations and a waste of information. Some models and data collections have more valuable results than others. Since this empirical research is the scoop in the area of intermodal transport, we also mention the difficulties and the unexpected outcomes. As already mentioned, the results could lead to a starting point of quantative conclusions and future policy actions.

In addition, we will discuss the outcome of the market share models, fully based on *Eurostat* data.

5.1 Model 1: General

A small recap: An individual is specified as an origin-destination, distinguished by its modality (Inland water, rail or sea). The dependent variable is the number of week departures of a certain individual.

The first model is fully based on *Ecorys* data. We will investigate the relations of the number of departures (frequency) and the regressors time, weekend percentage, modality sea and modality inland water. The underlying data is from 2015 to 2017. For each time frame there are 2401 individuals. In Table 5 parameter estimates together with their variances, t-values and p-values are presented. Where measured, we also present the robust estimates of the variance and their corresponding t-values and p-values.

The parameter estimate of the time-component addresses the relationship of time and frequency. In general, it is expected that time-consuming trips would have a lower frequency. For the fixed effects models, there is a very small positive relation, rejecting the expectation. The p-values say the estimates are not significant. For the random effects models, there is a small negative relation, confirming the expectation. Since it is known that the variable time changes only little through time, we expected a small parameter estimate. Besides, in the data set are relatively a lot of observations with a small frequency and a small time-value, which contradicts the found relationship. Moreover this freight transport database does not distinguish in type of freight. You can imagine that for certain types of products, e.g. flowers and food, the time and frequency relation would be more significant. At first sight, this parameter estimate does not lead to new and important insights. In a later model we only select origin destinations with a large frequency. Also, we will address the time variable per modality. This might lead to other results.

Next, we discuss the weekend-parameter. It is known that, due to (more conservative) legislation, employees are more expensive in the weekends. Therefore we expect that (if possible) most journeys would be executed during weekdays, whereby the percentage of weekend trips increases together with an increase in the frequency. Low frequent trips will mostly not be carried out in the weekends, whereas high frequent trips should be carried out in the weekends, for the reason of meeting the demand. According to this theory, the parameter of this variable is positive in all models, confirming our theory. Although we need to mention that the robust p-values mostly lead to rejections, which means that the relationship is not significant. This might be caused by the fact that other, more important factors might contaminate the relationship. That is why we only include this variable in the first model.

Next, we discuss the results for the time-invariant dummy variables for inland water and rail transport. Those are set out against the individuals that are transported by the sea. We expect that the frequency of certain origin-destinations by train and inland waterways is higher, as in

general distances are shorter and less freight volume could be transported. In all random effects models this parameter has a positive value. Also they are (very) significant. This is a first glimpse on the different relationships that the different modalities have with the the number of departures.

Note that the parameter estimates for the fixed effects Poisson model and the negative binomial model are almost equal. They differ in their (co)variances and t- and p-values. The overdispersion parameter of the negative binomial fixed effects model is 1.41. Though, a higher value of this parameter was expected, it still confirms the overdispersion.

Since the likelihood is specified differently for all models, we can only compare values for the same model cases, across different data collections. The (other) performance measures show that the fixed Poisson and negative binomial models perform best. Here, the improvement of specific count models to the more common linear models is confirmed. Also, note that the performance of the random effects models are better than OLS.

In the Appendix A.11 we give a clear overview of the effects for the different models. Obviously, the density of the estimated fixed effects is comparable for the Linear, Poisson and Negative binomial models. Differences in the fixed effects figures are hardly visible, since some fixed effects are (relatively) extremely high. For the random effects models, the distributions are obviously different.

We address the random effects versus the fixed effects by the Hausman test. This test gives 3.67 with a p-value of 0.158. This would not lead to rejection of the hypothesis that there are random effects. For feasibility reasons the Hausman test is only executed to the linear models, expecting that the Poisson and negative binomial models would lead to the same conclusions. This test gives an unexpected outcome, since the performance measures show that the fixed effects models have better fits.

Table 5: Parameter estimates of model 1[a].

	Linear			**Poisson**		**Negative binomial**	
	OLS	FE	RE**	FE	RE	FE	RE
(intercept)	3.627	($\mu = 6.927$, $\sigma^2 = 89.121$)	$\sim iid(2.731, 11.406)$	($\mu = 6.831$, $\sigma^2 = 89.765$)	$\sim Gamma(1.48, 0.50)$	($\mu = 6.932$, $\sigma^2 = 89.121$)	$\sim Beta(3.261, 2.401)$
time	-0.121	0.015	-0.001	0.010	-0.023	0.010	0.091
σ^2(Robust)	0.001	0.000 (0.000)	0.004	0.000 (0.000)	0.000 (0.000)	0.000	
t-value(Robust)	-4.751	2.491 (0.874)	-0.200	7.371 (0.731)	-22.161 (-8.423)	0.810	
p-value(Robust)	0.000	0.013(0.385)	0.841	0.000 (0.465)	0.000(0.000)	0.421	
weekend	0.021	0.303	0.269	0.112	0.116	0.112	0.601
σ^2(Robust)	0.141	0.004 (0.022)	0.049	0.000(0.024)	0.000 (0.003)	0.009	
t-value(Robust)	0.053	4.931(2.001)	1.221	7.130(0.712)	9.311 (2.071)	1.121	
p-value(Robust)	0.964	0.000(0.052)	0.222	0.000(0.474)	0.000(0.050)	0.261	
inlandW	10.970		11.601		1.621		2.401
σ^2(Robust)	0.023		0.627		0.000 (0.005)		
t-value(Robust)	28.131		14.642		277.020 (23.671)		
p-value(Robust)	0.000		0.000		0.000(0.000)		
rail	8.216		8.901		1.382		2.510
σ^2(Robust)	0.082		0.539		0.000(0.000)		
t-value(Robust)	28.662		16.490		352.174 (47.532)		
p-value(Robust)	0.000		0.000		0.000(0.000)		
MSD	143.8	11.1	55.4	11.1	52.4	11.1	51.4
MAD	5.473	1.156	2.101	1.151	2.280	1.151	2.271
MAPE	1.268	0.195	0.372	0.192	0.358	0.192	0.343
Likelihood*	-46591	-30301	-50531	-20999	-88898	-20999	-27087

*Log likelihoods are different for different models and can therefore not be compared, only models with the same likelihood functions can be compared.
**Via an R function the swar-transformation produced robust estimates of its (co)variance.

[a] Here the transport frequency is regressed on time, the percentage of weekend departures, a dummy for inland water and rail transport. The used data is from 2015 to 2017_1. Given are the coefficient estimate, its variance, t-value and its corresponding p-value. Where possible both the robust and the non-robust estimations are given.

5.2 Model 2: High frequency

For more insights in line with A) Mode differences and B) Socio-economic, financial and logistical factors, we present a model based on more frequently executed origin-destinations, or in other words based on origin-destinations with a higher number of departures. Data from 2015 to 2017 is used. The frequency treshold is arbitrarily chosen to be 10, meaning that only observations with a frequency higher than 10 will be included. This restriction leads to only 360 observations per half year. In the model the time-frequency relation will be mentioned per modality, which comes down to three different parameter estimates that encompass the time dependency apart from each other. Additionally, this model includes the *GDP per capita* variable and the *Quality of the port* variable (almost time-variant, so only present in random effects models) from the World Bank. Results for the parameter estimates are shown in Table 6.

Note again that a negative relation between time and frequency is expected, for all modalities. The inland water-time parameter estimate is controversial, since it has fluctuating estimates. This variable is not significant, which is partly due to the small amounts of observations of the inland water transport mode. The rail- and sea-time variables show a more reasonable outcome. The relationships are negative.

In more detail, for the linear model (fixed effects) an increase in time leads to a decrease of 2.05 in its number of departures (frequency). For the Poisson fixed effects model (and even so for the negative binomial model) the parameter estimate is -0.04. Assuming $\exp(\mathbf{x}'_{it}\boldsymbol{\beta}) = 1$, the marginal effect is equal to -0.04. Easier interpretable is the elasticity, which is -0.04 if the time variable equals 1. It shows that the time variable is inelastic. This means that the frequency of an origin-destination changes less than proportionally compared to the change in time an origin-destination journey takes.

Again, it is clear that the dependency of the change in the intermodal frequency is minimally explained by the time an origin-destination journey takes. Though it is significantly present for the origin-destinations with a high frequency. A comparable effect is measured for the sea transport. For now, we mentioned and scrutinized this relationship sufficiently.

Next, this model also leads to the first results for the so-called socio-economic, financial and logistical factors.

First we will analyze the parameter showing the relationship between the frequency and the quality of the ports (on country level). Obviously, this relationship is positive. It means that the port quality has a positive effect on the amount of intermodal transport. This is according to our expectation. Since the ports function as so-called transportation hubs, ports with a high quality would lead to more frequent rail transport as well. Therefore, the result brings quantitative support for policy on promoting intermodal transport.

It follows that the parameter of GDP per capita is significantly negative for all models. On first sight a negative relationship between the level of GDP per capita and the frequency of origin-destinations sounds unexpected. Though from an expert opinion this result is reasonable since countries with a higher GDP per capita in general also have a higher level of concentration in intermodal hubs, while countries with a lower GDP per capita in general have less hubs that facilitate the intermodal transport. In other words, the larger the GDP per capita the larger the probability of relatively large fragmentation of an intermodal network.

Next, the performance measures show (again) that the Poisson and the negative binomial fixed effects models perform best. A large Hausman-value with high significance leads to the conclusion that a fixed effects model fits the data best. An overdispersion parameter of 6.28 for the negative binomial fixed effects model confirms a high expected variance in comparison to the mean of the dependent variable.

Remarkably, again the fixed effects and random effects estimates fulfill for a large part in explaining the data as they are relatively large. This means that there might be a lot of latent information which could be useful for the model.

Table 6: Parameter estimates of model 2[a].

	Linear			**Poisson**		**Negative binomial**	
	OLS	FE	RE	FE	RE	FE	RE
(intercept)	7.14	($\mu = 41.55$, $\sigma^2 = 527$)	$\sim iid(5.37, 468)$	($\mu = 46.21$, $\sigma^2 = 1470$)	$\sim Gamma(5.00, 5.00)$	($\mu = 46.21$, $\sigma^2 = 1470$)	$\sim Beta(6.56, 2.76)$
inlandW/time	-3.08	1.27	-0.95	0.14	0.17	0.13	0.01
p-value(Robust)	0.00	0.28	0.28	0.49	0.09	0.58	
rail/time	-4.47	-2.05	-2.27	-0.04	0.31	-0.04	0.01
p-value(Robust)	0.00	0.04	0.02	0.00	0.00	0.00	
sea/time	-3.82	-1.26	-1.80	-0.09	2.39e-04	-0.10	-3.53
p-value(Robust)	0.00	0.00	0.00	0.00	0.56	0.00	
GDP per capita($\times 10^{-4}$)	-1.63	-2.98	-2.65	-0.12	-0.13	-0.12	-20.12
p-value(Robust)	0.00	0.00	0.00	0.00	0.01	0.00	
port quality	0.59		0.63		0.06		0.06
p-value(Robust)	0.00		0.00		0.00		
MSD	539.11	57.72	176.73	54.71	165.34	54.82	164.62
MAD	14.90	4.35	6.58	4.19	7.53	4.19	6.97
MAPE	0.78	0.23	0.32	0.21	0.41	0.21	0.35
Likelihood	-8167	-6173	-9037	-5682	-25744	-5724	-14848

[a] The transport frequency is regressed on time and transport mode variables, GDP per capita and the quality of the port. The data set is based on all observations from 2015-2017$_1$ with a frequency larger than 10. Given is the coefficient estimate and its corresponding p-value.

5.3 Model 3 & 4: Socio-economic and financial factors

Next, we add even more socio-economic and financial regressors into the model. In addition to the variables *GDP per capita* and *Quality of the port*, we add the variables *Export volume index*, *Population*, *Pump price Diesel* and *Km rail-lines*. Furthermore, we add a dummy variable for inland water and rail. The data set consists of the first period of 2015 and the first period of 2016. This is because for some (World Bank) regressors only data from the 2015 and 2014 can be used (where data from 2015 is connected to the *Ecorys* data set of the first period of 2016). Since there are relatively a lot of variables to estimate, estimation for the negative binomial random effects model seems to be very sensitive to its starting values and is infeasible. The results are shown in Table 7.

As in the first model, the time parameter concerning all modalities is fluctuating from positive to negative and has negative p-values, rejecting its significance.

The parameter concerning the export volume index shows in most models a positive relation, which at first sight seems reasonable. The frequency of intermodal transport would increase by an increase in the export volume. Exceptionally, this parameter has a very high p-value for the Poisson fixed effects model, only in the robust case.

Again the parameter for the relationship between the GDP per capita and the frequency is included. Remarkably, in the common OLS and the random effects model this parameter is positive in contrast to the other models. In these two cases the parameter has a very high p-value and will be rejected.

A next interesting variable is the population of a certain country. As expected, a larger population leads to more frequent intermodal transport. All models, with exception of the random effects Poisson model, show positive parameter estimates that are significant in all fixed effects cases. It seems like the random effects Poisson model is not functioning very well for the included variables. For the fixed effects models the elasticity for e.g. Germany is 0.6, which means that 10% increase in the population would lead to 6% increase in the number of departures. This sounds quite reasonable.

Although the random effects models do not function very well for this specific model, the positive relation between port quality and the amount of departures, as well as the positive values for the the inland water and rail parameters, are again present and significant.

Furthermore, the (almost) time-invariant parameters for diesel price and the total length of rail-lines in a country are not significant. This might be explained by the fact that the different modalities have different relationships with those variables. By a high Diesel pump price less road transport leads to more rail transport, but on the contrary a high Diesel pump price leads to less inland water transport. The amount of rail-lines might have a positive effect for the rail mode, but in the other way might lead to less other intermodal transport (and road transport). This relationship is not clear and significant in this model.

Again the Hausman test and the performance measures advocate the fixed effects instead the random effects. The overdispersion parameter of the negative binomial fixed effects model is 5.85.

Next, we estimate exactly the same model only including observations of the rail modality. Results are shown in Table 8. We will only discuss the parameter values for *Diesel pump price* and *KM rail-lines*. Since the effects of different modalities are removed, those parameter estimates are easier to interpret.

The parameter for diesel pump price is positive and significant. This is a logical and interesting result, since higher Diesel prices lead to relatively higher costs for road transport and inland waterway transport in comparison to the costs of rail transport. Thus, a higher diesel price might facilitate more frequent train departures (and a mode shift from road to rail transport).

Also a positive parameter value of the total amount of rail-lines is logical, since more rail-lines mean that the rail facilitation in a country is better and therefore lead to a higher amount of train departures.

Table 7: Parameter estimates of model 3[a].

	Linear			**Poisson**		**Negative binomial**	
	OLS	FE	RE	FE	RE	FE	RE***
(intercept)	-12.45	($\mu = -16.55$, $\sigma^2 = 35.23$)	$\sim iid(-10.39, 108.12)$	($\mu = 1.87$, $\sigma^2 = 27.67$)	$\sim Gamma(4.92, 4.91)$	($\mu = 1.86$, $\sigma^2 = 27.65$)	
time	-0.08	0.01	-0.03	1.33e-04	0.094	4.96e-04	
p-value	0.029	0.46	0.32	0.98	0.15	0.96	
export volume index	-0.01	0.04	0.01	8.15e-05	9.98e-05	8.42e-05	
p-value	0.00	0.00	0.08	0.94**	0.00	0.87	
GDP per capita($\times 10^{-5}$)	0.79	-5.34	-3.13	-1.08	0.16	-1.08	
p-value	0.50	0.00	0.00	0.00	0.93	0.00	
population($\times 10^{-7}$)	0.27	4.53	0.10	0.87	-0.03	0.87	
p-value	0.01	0.00	0.47	0.00	0.00	0.00	
port quality*	0.20		0.23		0.02		
p-value	0.00		0.00		0.00		
pump price Diesel*	0.01		0.00		0.00		
p-value	0.14		0.43		0.60		
km rail lines($\times 10^{-5}$)*	-2.23		0.25		1.03		
p-value	0.42		0.94		0.50		
inlandW	8.87		8.99		-0.00		
p-value	0.00		0.00		0.00		
rail	8.34		8.64		0.38		
p-value	0.00		0.00		0.00		
MSD	125.29	8.50	33.21	7.73	115.61	7.62	
MAD	5.01	0.98	1.64	0.92	4.44	0.92	
MAPE	1.33	0.23	0.35	0.16	1.65	0.16	
Likelihood	-21593	-14202	-23816	-9239	-47166	-9375	

*Since most observations are time-invariant (or almost), fixed effects estimation would be problematic.**Remarkably the robust p-value is high, while the non-robust p-value is 0.000.***Since there are relatively a lot of variables to estimate, estimation for the negative binomial random effects model seems to be very sensitve to its starting values and is infeasible for this case.

[a] The transport frequency is regressed on time, export volume, GDP per capita, population, the quality of the port, pump price of Diesel, number of km rail lines and a dummy for inland water and rail transport. Note that this is a data set for 2015-2016$_1$, all modalities included. Given is the coefficient estimate and its corresponding p-value.

Table 8: Parameter estimates of model 4[a].

	Linear			**Poisson**		**Negative binomial**	
	OLS	FE	RE	FE	RE	FE	RE
(intercept)	-18	($\mu = 27$, $\sigma^2 = 498$)	$\sim iid(-16, 148)$	($\mu = 663$, $\sigma^2 = 1740850$)	$\sim Gamma(5, 5)$	($\mu = 663$, $\sigma^2 = 1740850$)	
time	-0.32	-0.34	-0.48	1.08e-04	0.44	9.27e-05	0.05
p-value	0.39	0.43	0.24	0.00	0.00	0.00	
export volume index	0.02	0.12	1.00e-03	8.45e-05	1.01e-04	9.48e-05	1.00e-03
p-value	0.06	0.00	0.44	0.00	0.00	0.00	
GDP per capita($\times 10^{-5}$)	2.07	18.31	-7.86	-2.00	0.31	-2.00	30.00
p-value	0.05	0.00	0.01	0.00	0.91	0.00	
population($\times 10^{-8}$)	-7.13	-47.81	-11.82	-4.80	0.29	-4.81	0.10
p-value	0.13	0.01	0.03	0.00	0.86	0.00	
port quality	0.21		0.22		0.03		0.04
p-value	0.00		0.00		0.00		
pump price Diesel	0.07		0.08		3.68e-07		1.23e-05
p-value	0.02		0.02		0.00		
km rail lines($\times 10^{-5}$)	33.61		41.62		0.22		0.01
p-value	0.00		0.00		0.33		
MSD	187.51	19.93	48.92	16.84	101.41	16.80	85.11
MAD	8.14	1.84	2.84	1.67	4.88	1.67	3.66
MAPE	1.37	0.22	0.41	0.17	0.73	0.17	0.65
Likelihood	-7276	-5302	-8009	-3887	-17729	-3962	-4058

[a] The transport frequency is regressed on time, export volume, GDP per capita, population, the quality of the port, pump price of Diesel and the number of km rail lines. Note that this is a data set for 2015-2016$_1$, taken only rail data. Given is the coefficient estimate and its corresponding p-value.

5.4 Model 5: General forecasts

Comparably to the now mentioned models, forecasts will be made for 2016 based on parameter values using data from 2014 to 2015. We include the variables *GDP per capita* and *Quality of the port*, and the combination of time and the three modalities. We estimate the model based on 813 observations. The results are presented in Table 9.

Clearly the performance measures again advocate fixed effects models. The linear fixed effects model (unfortunately) does a better job in forecasting than the Poisson and negative binomial models, based on the MSD (and MAD). On the contrary, the Poisson and negative binomial models do a better job in forecasting than the linear models based on the MAPE.

5.5 Model 6 & 7: Geographical differences

Next we should address the geographical differences of intermodal transport. First this will be done at country level. Therefore we make use of the cluster random and fixed effects models. Origin-destinations will simply be assigned to the country cluster. Furthermore in the analysis we added the variables *GDP per capita*, *Quality of the port* and the combination of time and the three modalities. In addition we added also dummy variables for the modalities rail and sea. Estimates are based on the data set from 2015 to 2016. The results are shown in Table 10.

We will not discuss the parameter estimates thoroughly, as they mostly agree with the already mentioned models. Only the fixed and random effects, which are based on the belonging country, will be addressed.

Remarkably, the linear random effects model performs better than the linear fixed effects model according to its MSD, MAD and MAPE. This suggests that the latent part for country clusters could be randomly explained, meaning that the number of departures fluctuates within the countries. For the Poisson models, the fixed effects model performs way better and performs better than the linear random effects model according to its MSD and MAPE.

In Figure 6 a visualization of the Poisson fixed effects is presented on country level. Here, darker areas have higher fixed effects (white areas are not taken into account). It is clear that in the Netherlands, as it know looks like the epicentrum of intermodal transport (what a proud!), the amount of departures at specific origin-destinations is highly determined by a latent part. Then Germany and France follow. And since the colour fades away from the Netherlands, the latent part in the model becomes less, or in theoretical terms: the country clustered fixed effects are lower. For a large part this could be explained since the Netherlands (together with France and Germany) is literally a central point in Europe, where several transport routes should cross. Products from Italy will first be transported to the Netherlands, France or Germany before transport to Scandinavian countries. This seems to be a logical explanation of this outcome.

Next, we execute the cluster analysis at TEN-T level. The TEN-T Programme was established by the European Commission to support the construction and upgrade of transport infrastructure across the European Union (European Commission, 2016). This European program aims at certain TEN-T routes, which means that the intermodal transport development and the increase of transport at those routes needs to be promoted. We calculate the models based on a cluster of the origins belonging to a certain TEN-T against a cluster of all other origins. This means that for the interesting TEN-T routes different models will be made (since some origins do overlap in the different TEN-T's). The clustered fixed effects from both a linear and Poisson model are presented in Table 11.

A TEN-T route that already functioned according to its aims (since it concerns an industry-rich area), is the Rhine-Alpine TEN-T. As we see, the number of departures at this route are for a significantly larger part explained by latent variables, leading to a higher number of departures. We find this a reasonable outcome since in the basic model a certain TEN-T or river variable or industry area is not accounted for.

Contrary to the Rhine-Alpine route, for the Baltic Adriatic TEN-T the number of departures at this route are less explained by a latent part, leading to a relatively smaller number of departures in this area in comparison to the overall latent part. We find this intuitively reasonable since economic relations in Europe are mostly from the east to the west.

Furthermore for the North Sea Baltic, the number of departures at this route are for a significantly larger part explained by a latent part. There is a large stream of intermodal water transport (sea and inland waterways) in and around the North Sea.

Lastly, for the Mediterranean route the number of departures at this route are less explained by a latent part. This is simply caused by the fact there is a low amount of infrastructure in these areas.

All in all, again some logical and interesting quantitative insights are given through the fixed effects model specification.

Table 9: Parameter estimates and forecasts of model 5[a].

		Linear			Poisson		Negative binomial	
		OLS	FE	RE	FE	RE	FE	RE
Parameters	**(intercept)**	-6.16	($\mu = 7.29$, $\sigma^2 = 27$)	$\sim iid(1.75, 244.10)$	($\mu = 6.23$, $\sigma^2 = 105.23$)	$\sim Gamma(4.43, 5.14)$	($\mu = 6.23$, $\sigma^2 = 105.40$)	$\sim Beta(6.56, 2.76)$
	inlandW/time	-0.30	-2.31	-0.24	-0.20	0.25	-0.20	-0.01
	rail/time	0.77	2.04	0.80	-0.15	0.35	-0.11	-0.07
	sea/time	-0.75	-4.56	-0.71	-0.10	-0.09	-0.10	-0.21
	GDP per capita($\times 10^{-5}$)	-6.63	5.29	-2.68	-1.37	-0.23	-1.37	40.0
	port quality	0.34		0.17		0.04		0.05
In-sample performance measures	MSD	250.14	2.96	54.52	2.86	105.92	2.94	97.80
	MAD	7.63	0.52	2.21	0.51	3.82	0.58	3.33
	MAPE	2.03	0.08	0.50	0.06	0.82	0.09	0.75
	Likelihood	-6554	-3155	-7248	-2661	-13544	-5721	-6472
Performance measures forecast 2016	MSD	405.43	94.58	156.87	102.59	180.31	102.21	138.88
	MAD	9.61	3.68	4.87	3.76	6.49	3.73	6.06
	MAPE	2.09	0.46	0.78	0.34	0.97	0.33	0.90

[a] The transport frequency is regressed on time and transport mode variables, GDP per capita and the quality of the port. The data set is based on all observations from 2014-2015. Forecasts are made for 2016_1. Given are its performance measures, both in-sample and out-of-sample.

Table 10: Parameter estimates of model 6[a].

	Linear		**Poisson**	
	FE	RE	FE	RE
(clustered effect)	($\mu = 10.59$, $\sigma^2 = 24.00$)	$\sim iid(0.42, 134.22)$	($\mu = 10.21$, $\sigma^2 = 0.69$)	$\sim Gamma(1.19, 2.01)$
inlandW/time	-1.49	-0.21	0.19	0.48
p-value(Robust)	0.00	0.70	0.00	0.00
rail/time	-0.19	-1.54	0.19	-0.17
p-value(Robust)	0.10	0.00	0.00	0.01
sea/time	-0.04	-0.01	-0.13	-0.01
p-value(Robust)	0.00	0.94	0.00	0.14
GDP per capita($\times 10^{-5}$)	-8.18	-0.46	0.64	2.44
p-value(Robust)	0.00	0.00	0.00	0.00
port quality		0.15		0.00
p-value(Robust)		0.00		0.00
rail		2.62		1.58
p-value(Robust		0.08		0.00
sea		-5.31		0.77
p-value(Robust)		0.00		0.04
MSD	150.31	146.01	145.58	195.57
MAD	5.91	4.83	5.53	5.88
MAPE	1.87	0.99	1.53	0.67
Likelihood	-19121	-243923	-26639	-64096

[a] The transport frequency is regressed on time and transport mode variables, GDP per capita and the quality of the port (dataset of 2015-2016_1). Clusters are based on the origin country.

Table 11: Fixed effects of model 7[a].

	Linear fixed effect	**Poisson fixed effect**
Rhine-Alpine	15.08	17.16
Others	4.95	7.56
Baltic-Adriatic	4.07	5.29
Others	5.97	8.77
North Sea Baltic	11.91	12.57
Others	5.61	8.19
Meditarrenean	2.23	3.76
Others	6.07	8.95

[a] The fixed effects for the TEN-T corridors are presented. Those come from a model where transport frequency is regressed on time and transport mode variables, GDP per capita and the quality of the port (data set of 2015-2016_1).

5.6 Case study: Market share model

At last, we discuss the results of the case study on aggregated data. A market share model is introduced. For rail, road and inland water the counrty shares of are calculated and used as the dependent variables in the model. The shares are given in the Appendix A.10 Figure 9 (for comparison reasons together with the country shares of other variables). According to the method section, the United Kingdom will function as benchmark country. The results for the market share model are shown in Table 12.

At first it becomes clear that for all modalities the outcomes of the two approaches are different. The parameter estimates are relatively comparable in their size and sign. For the rail and inland water market shares models the estimates are very comparable. There is a large difference in the (co)variances of the approaches, leading to different p-values. Furthermore there is a difference in the MSD and likelihood. The OLS-estimator by definition leads to the smallest MSD. On the contrary this approach leads to the smallest log likelihood and the largest p-values.The SUR estimator leads to higher likelihoods. As Fok, Franses and Paap (2002) state that the SUR-estimates maximize the log likelihood function, we will discuss the results of this estimator.

We discuss the elasticities. The SUR-parameter estimate for GDP has an elasticity of $(1 - M_{j,t}) *$ 0.15, which comes down to an elasticity of 0.12 for Germany and 0.14 for the Netherlands (taking the market shares of 2013). This means that a 1% increase in GDP leads to respectively 0.12% and 0.14% increase in the road market share. In case of the rail shares those elasticities are respectively 0.77 and 1.19. In case of waterway shares this is 0.27 for Germany (The amount of freight in tonne-km for the Netherlands is unknown). These positive relationships between the market shares and the GDP seem reasonable (and significant).

Furthermore, the relationship between population and market shares for all modalities is positive and therefore reasonable as well.

Since the environment tax and policy climate index change less through the years and furthermore differ (relatively) less, it is expected that they will function as the overall constant variable[7].

There seems to be a negative relationship between the market shares of rail and road, and the trade and infrastructure investment variables. On the contrary this relationship is positive for the road transport, which is a more reasonable result.

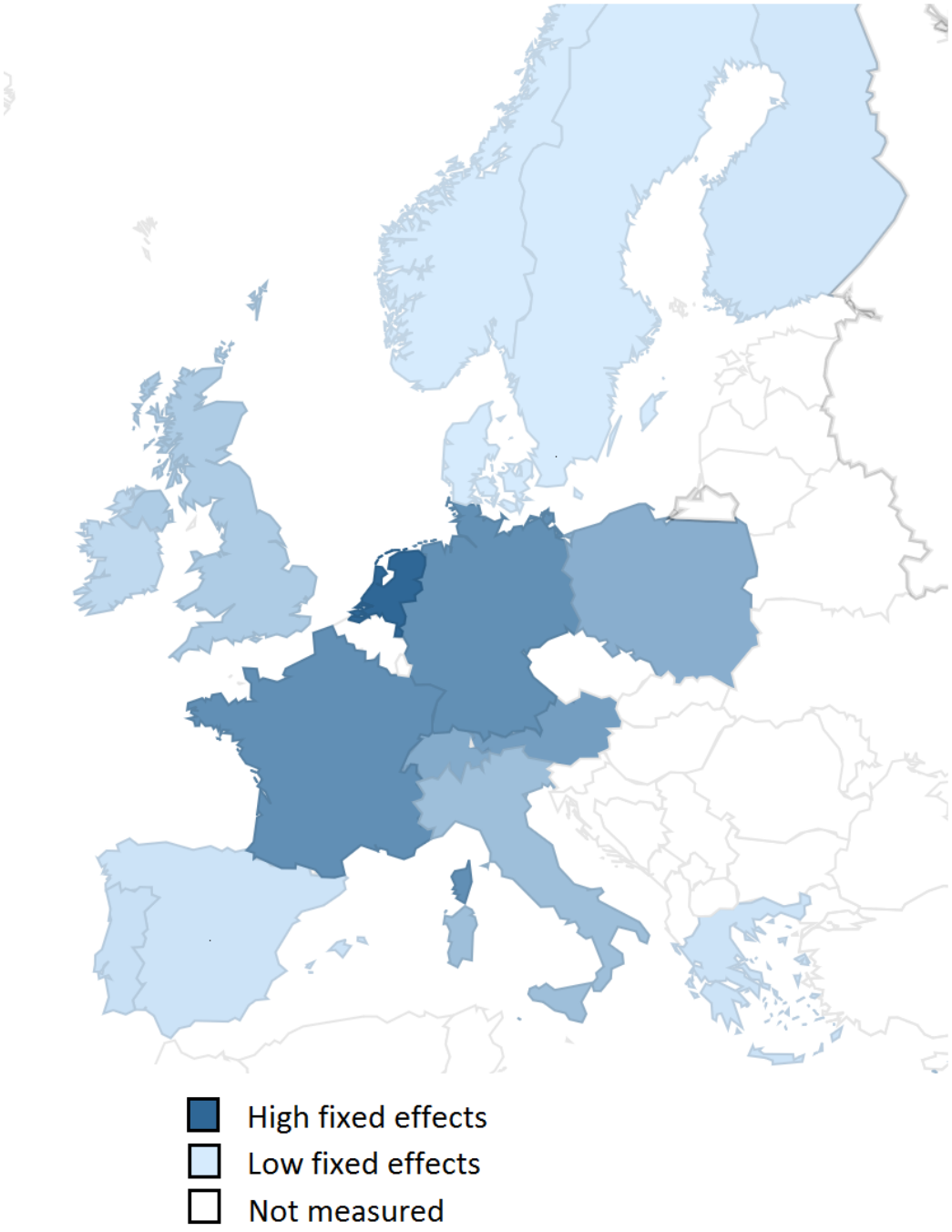
Important factors still seem to be GDP and population. Unfortunately, the results of the market shares do not lead to other insights. This might partly be caused by the few amount of observations that are available. Also, this shows again the complexity of the relationships between certain variables, meaning that small changes in for example environment tax do not directly lead to significant changes in the market shares. Furthermore, it shows the added value of the data at a disaggregated level and the results from the random and fixed effects model analysis.

[7]The complete analysis is also executed including a constant. Unfortunately this leads to insignificant parameter estimates (only the constant is significant).

Table 12: Parameter estimates for the market share models.

	Road		**Rail**		**Inland water**	
	OLS	SUR	OLS	SUR	OLS	SUR
GDP	2.08	0.15	0.63	1.21	0.68	0.92
p-value	0.018	0.05	0.59	0.00	0.68	0.05
Population	1.14	0.67	0.41	0.52	0.46	0.59
p-value	0.03	0.00	0.49	0.000	0.49	0.00
Trans.Inv.	-0.17	0.05	-0.04	-0.21	-0.09	-0.11
p-value	0.67	0.00	0.93	0.00	0.88	0.28
Trade	-0.63	0.05	-0.16	-0.06	-0.16	-0.19
p-value	0.07	0.00	0.57	0.00	0.69	0.00
Env.Tax.	0.19	0.39	0.57	0.03	1.11	0.39
p-value	0.85	0.00	0.55	0.00	0.52	0.20
Policy	0.13	0.02	0.05	-0.07	0.05	0.16
p-value	0.75	0.05	0.05	0.00	0.91	0.03
MSD	1.79	2.35	1.78	1.97	2.12	2.18
Likelihood	-666	-544	-627	-542	-450	-427

Figure 6: This figure shows the size of the fixed effects for European countries relative to each other. It is based on the model where transport frequency is regressed on time and transport mode variables, GDP per capita and the quality of the port. A dataset of 2015-2016$_1$ is used. The countries form separate clusters.



6 Conclusion

In this paper we used both linear and count fixed and random effects models to analyze the European intermodal transport flows. The models provide us with attractive model representations used for the the estimation of fixed and random effects parameters, as well as common parameters. In addition, we used a market share model to analyze causal relations explaining changing country shares for certain modalities.

Governments and intermodal operators are in need of quantification of the European intermodal transport flows. Until now there is limited econometric research. This paper acts as a starting point. The research answers the following main question: "How was European intermodal freight transport organised in the past years and what factors promote a modal shift?" For the sake of clarity, intermodal trends are assessed along the following three lines: A) Mode differences, B) Socio-economic, financial and logistical factors, and C) Geographical differences. The mentioned models are used to conduct analysis that lead to interesting quantified insights.

Clearly, the models show that rail and inland waterway transport generally have a relatively larger number of departures than sea transport. Whereas the relationship between time and the number of departures for a certain origin-destination is not significant, models based on high frequent departures show a negative relationship between time and the number of departures, for each modality. This means that longer journeys depart less frequently. Next, it becomes clear that when the GDP per capita is higher, departures for origin-destinations are less frequent. This finding confirms that those countries generally have a more fragmented transport network. In addition, countries with a higher GDP have a larger share of transport, for all modalities. Thus, in countries with a higher GDP per capita intermodal transport is more fragmented, but the total (intermodal) transport is relatively higher. Further, we concluded that countries with a higher export volume have more frequent departures at specific origin-destinations. Another interesting finding is that high port quality leads to an increase in intermodal transport. Based on this finding policy on promoting a modal shift should focus on improvement of the ports. Furthermore, higher amount of rail lines and a higher pump price logically lead to a higher amount of train departures. The pump price could be taxed to increase the amount of departures by train. Also, from the clustered fixed effects models it becomes clear that the Netherlands, together with Germany and France, function as a central intermodal transport spot. Clearly, based on the height of the latent part (the fixed effects) the amount of departures lays higher for the Rhine-Alpine and North Sea Baltic TEN-T corridors, whereas it lays lower for the Baltic Adriatic and Mediterranean TEN-T corridors. Those first results indicate that the Rhine-Alpine and North Sea Baltic TEN-T corridors are promoted with success.

We stated two methodological questions to critically asses the validity of the data and the models.

Overall, it is clear that the models are valuable for quantification of the intermodal transport flows. Conclusions mainly based on qualitative research could from now on be drawn and quantified by the fixed and random effects models. Note that the fixed effects models functioned clearly better at the origin-destination level, since the latent part in the number of departures is not random, but explained by latent (omitted) variables correlated to the present variables. However, for the clustered approach, the random effects models approximate the fixed effects models' performances. Both models are valuable. Overall the Poisson and negative binomial models perform better than the linear models.

Furthermore it is clear that *Ecorys* data at origin-destination level, in other words data at disaggregated level, add value in the research. Only aggregated data (country level) hardly lead to significant quantifications by a lack of observations. Therefore, collection of the data on origin-destination level should be continued.

It is widely known that road transport still remains dominant in the transport world. But since a modal shift could lead to an improvement of cost effectiveness, economic growth and the reduction of social and environmental externalities, the intermodal shift should be promoted. This research functions as a good starting point to promote the modal shift.

7 Discussion

In this section we will highlight difficulties and problems regarding the data, the used models, validity of the outcomes and we address some recommendations for further research.

At first, we need to mention that the data from *Ecorys* bring opportunities for quantification in the intermodal transport field (in comparison to the data on country level). Since the collection of this valuable asset, the data on origin-destination level, just started, a lot of improvements could be made. For example it is known that in the data there is still a lack on the amount of East-European origin-destinations (Poland and further East). Also it is clear that the data collection started from the Netherlands, since relatively a lot of departure spots are established in the Netherlands. For this reason, by including a geographically broader zone in the future the conclusions can be toughened. Furthermore, the data could be improved by adding other variables. Information about the transport distance, the freight volume, the transported products and the costs of transport will generate new insights. Those extensions cannot be made to the data on country level.

Quite interesting is the fact that the *Ecorys* data set does not contain any missing observations. This is because the collection process is set up bottom-up. Only the known and most basal information is included. Since missing data in a sense can be informative, stating that this leads to non-missing information/data on the side, a more top-down collection process could be started besides. This questions in what way the dataset should and could be improved.

A more specific point is the fact that the data set consists of many origin-destinations that are only executed once a week. In fact some of those origin-destinations are executed more frequently. However, the carriers that also provide this origin-destination journey are not included in the data set. For a better picture of the intermodal landscape, those should be included in the future. Lastly, about the data, a lot of information needed to be thrown away because the data sets are analyzed over time and needed to be aggregated. This removed (informative) data could still be analyzed.

Apparently, the calculation of robust errors is not always comparable between the models, since different specifications of the robust standard errors were given. Clearly, for all models this robust error lead to a higher (co)variance and therefore higher p-values. For the random effects models, the estimations did not always lead to the same conclusions. This might be encouraged by the infeasibility of the maximum likelihood to maximize over a relatively large amount of parameters. This is also visible in some results. For example, for some model specifications the gamma parameters did hardly change in comparison to its starting values. This is a common econometric problem, but should be mentioned. Improvement on this feasibility could be made.

Furthermore, negative binomial fixed and random effects models could be problematic. Several specifications of the negative binomial fixed effects models come down to the Poisson model. This is often mentioned as being problematic in the literature and needs further research.

Despite some differences of the outcomes in the fixed and random effects models, overall there were valid outcomes based on the significance of parameter estimates.

In contrast, the outcomes from the market share model were not very informative. The OLS-estimates were completely different from the SUR-estimates. Both outcomes are doubtful since there were too less observations.

Some outcomes from this research could function as a starting point for policies of the European Commission. Based on expert knowledge there are two findings that are mainly useful for certain policy actions. The first is the result that port quality leads to an increase in the number of intermodal transport departures. To promote intermodal transport Europe should invest in the port infrastructure. The second is the finding that tax on Diesel leads to an increase in transport by rail. This confirms the policy of certain European countries on increasing their Diesel tax for a better climate. Those are valid quantifications, but still need to be scrutinized further.

In general, still a lot of research should be done. It is quite clear that this intermodal research area needs quantification. For further research more econometric analysis could be generated. First of all, mixed models could be applied on the data since even if the effects are mainly fixed, a random effect still could be present. Furthermore, there might be options to transform qualitative knowledge to quantitative knowledge. A rather Bayesian approach might generate interesting results as well: it assigns a hypothetical distribution to the known data and thereby might enrich the research.

REFERENCES

Allison, P. D., Waterman, R. P. (2002). Fixed-Effects Negative Binomial Regression Models. Sociological methodology, 32(1).

Agamez-Aris, M. Moyano-Fuentes, J. (2017). Intermodal transport in freight distribtution: a literature review. Transport Reviews.

Abramowitz, M. Stegun, I.A. (1972). Handbook of mathematical functions and formulas, graphs and Mathematical tables. John Wiley and Sons. 900-924.

Blauwens, G. Vandaele, N. Van de Voorde, N. Vernimmen, B. Witlox, F. (2006). Towards a Modal Shift in Freight Transport? A Business Logistics Analysis of Some Policy Measures. Transport Reviews.

Bontekoning, Y.M. Macharis, C. Trip, J.J. (2003). Is a new applied transportation research field emerging? - A review of intermodal rail-truck freight transport literature. Elsevier.

Cameron, A. Colin., Trivedi, Pravin K. (2005). Microeconometrics: Methods and Applications. Cambridge University Press.

Cameron, A. Colin., Trivedi, Pravin K. (1998). Regression analysis of count data. Cambridge University Press. 275-300.

De Jong, G. Gunn, H. Walker, W. (2004). National and International Freight Transport Models: An Overview and Ideas for Future Development. Transport Reviews..

European Commission (2017). Amendment of the combined transport directive.
https://ec.europa.eu/transport/themes/urban/consultations.

European Commission (1997).Intermodality and intermodal freight transport in the European Union. Communication paper from the commission to the European parliament and the council.

European Commission (2016). European mobility week 2016: Sustainable transport is an investment for Europe. *https://ec.europa.eu/transport/media/news.*

European Commission (2016). TEN-T. *https://ec.europa.eu/inea/en/ten-t.*

Eurostats, Statistics explained (2017). Freight transport statistics-modal split.
http://ec.europa.eu/eurostat/statistics-explained/index.php

Eurostats (2017). Several data sets, further specified in the text, are used as variables in the models.
http://ec.europa.eu/eurostat/data/database

Fok, D. Franses, P.H. and Paap, R. (2002). Econometric analysis of the market share attraction model. In Advances in Econometrics (pp. 223-256). Emerald Group Publishing Limited.

Greene, W. (2001). Estimating econometric models with fixed effects.

Hausman, J. A., Hall, B. H., Griliches, Z. (1984). Econometric models for count data with an application to the patents-RD relationship. Econometrica 52 ,909-938.

Ismail, N. Jemain, A. A. (2007). Handling overdispersion with negative binomial and generalized Poisson regression models. In Casualty Actuarial Society Forum.(pp. 103-158).

Kreutzberger, E. Macharis, C. Vereecken, L. Woxenius, J. (2003). Is intermodal freight transport more environmentally friendly than all-road freight transport? A review. *Paper presented at the NECTAR Conference No 7, Sweden.*

OECD, (2001). Intermodal freight transport: institutional aspects.
http://www.oecd-ilibrary.org.eur.idm.oclc.org

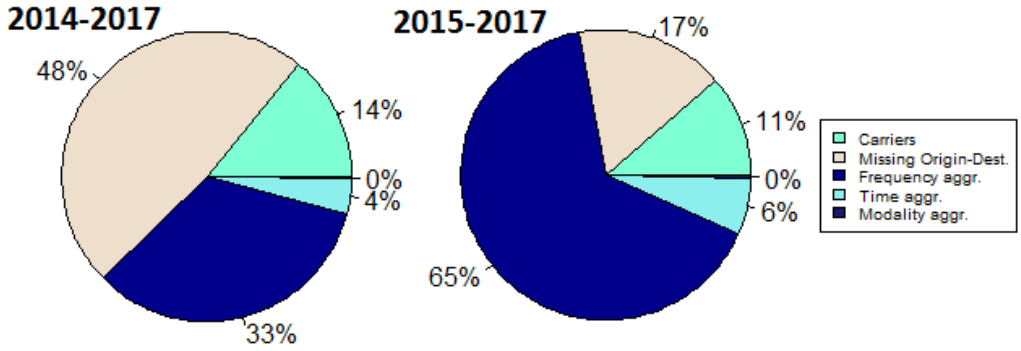
Schreiber, Sven. "The Hausman test statistic can be negative even asymptotically." Jahrbücher für Nationalökonomie und Statistik 228.4 (2008): 394-405.

A Appendix

A.1 Explanation of the aggregation process of the *Ecorys* data

Here we explain the aggregation process of the *Ecorys* data set. We have data sets per half year. We need to aggregate those data sets to create a panel data set. When aggregating the data sets there is a loss of observations, since not every observation is in all half years. We should aggregate at several levels. In Figure 7 the different levels of data aggregation are shown together with the percentage observation loss, both for aggregation of 2014 with 2017 and of 2015 with 2017. The order of aggregation is according to the order of the legend.

Figure 7: The percentage whereby each aggregation level shrinks the amount individuals, or in other words: the percentage each aggregation level adds to the total loss of observations. The order of aggregation is according to the order of the legend.



In the raw database the data are at carrier level. When aggregating the data set of 2014 with the data set of 2017, 14% of the total loss of observations in the data set is caused by removing carriers that are not in both years (In the dataset from 2017 more carriers are present). 48% of the total observation loss is caused by missing origin-destinations, mostly due to a comparable reason: the data set from 2014 contains less origin-destinations, even for the same carriers as in 2015. A next 33% is caused by aggregating the frequency of origin-destinations, counting up frequencies from different carriers for the same origin destinations (and the same modality). Only 4% is due to aggregation for different transport times (calculating a mean transport time in case of aggregation), which means that most origin-destination individuals have the same transport times (and if not they differ according to their transport mode). Clearly, a last aggregation over the transport mode encompasses a minimal percentage in loss of observations. This is due to the fact that certain origin-destinations are mostly concerned with certain modalities. For example, in Düsseldorf there is no sea, which makes transport by sea impossible. Moreover this last aggregation will be seen as a loss of information and will in general not be executed.

In Table 13 only aggregation per half year apart from other half years is given. Clearly, the number of observations increases during the (half) years - in exception of the first and second data set in 2015, that almost stays the same.

In Table 14 statistics are given for the aggregation of datasets over the years from 2014 to 2017. This does not take into account the earlier mentioned aggregation over the modalities, as this variable is informative in the model. Since the aggregation over all years leads to a large loss of observations, further analysis mostly refers to the data set aggregated from 2015 to 2017. The statistics of this aggregation are given in Table 2. As you can imagine aggregation over a selection of other half years leads to inclusion of different observations.

All in all, this preparation process raised the questions and complexities that aggregation could lead to. The trade-off between the loss of observations by the amount of years and the loss of observations by aggregation over the included years leads to different and interesting dynamics and conclusions in this research. Obviously, this should not be neglected.

Table 13: Statistics of the number of departures per half year, without aggregation over the years[a].

		2014	**2015$_1$**	**2015$_2$**	**2016$_1$**	**2016$_2$**	**2017$_1$**
All modalities	#obs	1516	4254	4247	4935	5009	5468
	mean	6.73	5.11	5.19	5.28	5.13	5.45
	median	3	2	2	2	2	2
	mode	1	1	1	1	1	1
	var	184.23	101.06	107.79	129.03	116.08	136.77
	max	188	209	188	188	190	236
	% Frequency = 1	30.47	42.29	41.53	40.97	39.50	38.89
Rail	#obs	380	1319	1309	1365	1363	1365
	mean	12.41	9.39	9.64	10.82	10.34	10.56
Inland water	#obs	248	323	329	349	365	368
	mean	12.30	11.96	11.87	12.49	13.01	13.21
Sea	obs	888	2612	2599	3221	3281	3735
	mean	2.75	2.10	2.12	2.15	2.08	2.81

[a] It means that all the carriers, modes and origin-destinations are selected, even if not present in all half years.

Table 14: Statistics of the number of departures per half year, completely aggregated over the five half years[a].

		2014	**2015$_1$**	**2015$_2$**	**2016$_1$**	**2016$_2$**	**2017$_1$**
All modalities	#obs	813	813	813	813	813	813
	mean	8.10	8.6	8.84	10.69	10.06	10.07
	median	4	4	4	4	4	4
	mode	1	1	1	1	1	1
	var	247.38	271.12	285.92	418.96	377.65	376.20
	max	188	209	188	188	188	188
	Frequency = 1	18.94	17.10	16.97	16.97	17.47	17.22
Rail	#obs	198	198	198	198	198	198
	mean	13.1	13.22	13.65	19.45	17.36	17.46
Inland water	obs	190	190	190	190	190	190
	mean	13.57	14.79	15.18	16.51	17.47	17.36
Sea	#obs	425	425	425	425	425	425
	mean	3.32	3.70	3.76	4.00	3.34	3.36

[a] It means that only the carriers, modes and origin-destinations are selected which are present in all half years.

A.2 General form robust sandwich (co)variance estimate

The general robust sandwich (co)variance estimate $\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{B}}\hat{\boldsymbol{A}}^{-1}$ is formed by taking the derivatives of the likelihood function $q(y, \boldsymbol{x}, \boldsymbol{\theta})$:

$$\boldsymbol{A} = E\Big[\frac{\delta^2 q(y, \boldsymbol{x}, \boldsymbol{\theta})}{\delta\boldsymbol{\theta}\delta\boldsymbol{\theta}'}\Big], \tag{32}$$

$$\boldsymbol{B} = E\Big[\frac{\delta q(y, \boldsymbol{x}, \boldsymbol{\theta})}{\delta\boldsymbol{\theta}}\frac{\delta q(y, \boldsymbol{x}, \boldsymbol{\theta})}{\delta\boldsymbol{\theta}'}\Big]. \tag{33}$$

A.3 Elaboration on the linear random effects model

For (3) we require estimates of $\hat{\sigma}_\epsilon^2$ and $\hat{\alpha}_\epsilon^2$ (Cameron, 2005). We calculate $\hat{\sigma}_\epsilon^2$ by the Within method:

$$\hat{\sigma}_\epsilon^2 = \frac{1}{N(T-1)-K}\sum_{i=1}^{N}\sum_{t=1}^{T}\big((y_{it}-\bar{y_i}) - (\boldsymbol{x}_{it}-\bar{\boldsymbol{x}}_i)'\boldsymbol{\beta}_W\big)^2. \tag{34}$$

From the Between method of $\bar{y_i}$ on an intercept and $\bar{x}_i$, we obtain the other variance component:

$$\hat{\sigma}_\alpha^2 = \frac{1}{N-K-1}\sum_{i=1}^{N}(\bar{y_i} - \hat{\mu}_B - \bar{\boldsymbol{x}}_i')\hat{\boldsymbol{\beta}}_B)^2 - \frac{1}{T}\hat{\sigma}_\epsilon^2. \tag{35}$$

A.4 General form of Gaussian Quadrature

Suppose the following integration needs to be computed (θ is given):

$$\int f(y_i|;x_i;\theta;\alpha_i)g(\alpha_i;\theta)\delta\alpha_i. \tag{36}$$

Numerical integration means making a grid of P points $\alpha_i^1 < \cdots < \alpha_i^P$ and evaluating the surface of the function between the grid points. By Gaussian Quadrature a weighted sum of the function values is calculated. Therefore equation (36) is rewritten and approximated as follows:

$$\begin{aligned}
&\int f(y_i|;x_i;\theta;_i)g(\alpha_i;\theta)\delta\alpha_i \\
&= \int w(\alpha_i)\frac{f(y_i|;x_i;\theta;\alpha_i)g(\alpha_i;\theta)}{w(\alpha_i)}\delta\alpha_i \\
&\approx \sum_{p=1}^{P} w_p \frac{f(y_i|;x_i;\theta;\alpha_p)g(\alpha_p;\theta)}{w(\alpha_p)}.
\end{aligned} \tag{37}$$

The function $w()$ is chosen according to the region of α_i; Values of w_p depend on the chosen function and can be found in the handbook of Abramowitz and Stegun (1972, page 900-924).

A.5 Result of the log likelihood of the Poisson random effects model

The Poisson model with gamma-distributed random effects density g($\alpha_i|\theta_1$,θ_2) results to the unconditional joint density for observation i:

$$f(\boldsymbol{y}_i|\boldsymbol{X}_i, \boldsymbol{\beta}, \boldsymbol{\theta}) = \int\prod_t \Big[\exp(-\alpha_i\lambda_{it})(\alpha_i\lambda_{it})^{y_{it}}/y_{it}!\Big]\theta_2^{\theta_1}\alpha_i^{\theta_1-1}exp(-\alpha_i\theta_2)/\Gamma(\theta_2)\delta\alpha_i. \tag{38}$$

Since an analytic result for this integral seems to be infeasible, the integral will be approximated by the numerical Gauss–Hermite quadrature method (see Appendix A.4 for a general explanation). We get the following integral and approximation:

$$
\begin{aligned}
f(\boldsymbol{y}_i|\boldsymbol{X}_i,\boldsymbol{\beta},\boldsymbol{\theta}) &= \int \prod_t \Big[\exp(-\alpha_i\lambda_{it})(\alpha_i\lambda_{it})^{y_{it}}/y_{it}!\Big]\theta_2^{\theta_1}\alpha_i^{\theta_1-1}\exp(-\alpha_i\theta_2)/\Gamma(\theta_2)\delta\alpha_i \\
&= \int \frac{w(\alpha_i)}{w(\alpha_i)} \prod_t \Big[\exp(-\alpha_i\lambda_{it})(\alpha_i\lambda_{it})^{y_{it}}/y_{it}!\Big]\theta_2^{\theta_1}\alpha_i^{\theta_1-1}\exp(-\alpha_i\theta_2)/\Gamma(\theta_2)\delta\alpha_i \\
&\approx \sum_{p=1}^{P} \frac{w_p*}{\exp(-\alpha_p)} \prod_t \Big[\exp(-\alpha_p\lambda_{it})(\alpha_p\lambda_{it})^{y_{it}}/y_{it}!\Big]\theta_2^{\theta_1}\alpha_p^{\theta_1-1}\exp(-\alpha_p\theta_2)/\Gamma(\theta_2).
\end{aligned} \tag{39}
$$

By definition of the gamma function, α_i could only be positive. Therefore the so-called Laguerre integration is followed. Here, $w(\alpha_p)$ is $\exp(-\alpha_p)$ and w_p* follows from the handbook of Abramowitz and Stegun (1972). The maximum log likelihood function of the Poisson random effects model, based on the Laguerre integration, is approximated as follows:

$$
\sum_{i=1}^{n} \log \sum_{p=1}^{P} \frac{w_p*}{\exp(-\alpha_p)} \prod_t \Big[\exp(-\alpha_p\lambda_{it})(\alpha_p\lambda_{it})^{y_{it}}/y_{it}!\Big]\theta_2^{\theta_1}\alpha_p^{\theta_1-1}\exp(-\alpha_p\theta_2)/\Gamma(\theta_2). \tag{40}
$$

In comparison to previous log likelihoods, the logarithm of this log likelihood cannot be taken further inside and transform log $\prod_t$ into $\sum_t$ log. As the $\prod_t$ could lead to infeasible calculations - in the data set this is the case - the formula should be rewritten taking an exponent and a log of the term inside the product over p:

$$
\sum_{i=1}^{n} \log \sum_{p=1}^{P} \exp\Big(\log(\frac{w_p*}{\exp(-\alpha_p)}) + \log(\prod_t \Big[\exp(-\alpha_p\lambda_{it})(\alpha_p\lambda_{it})^{y_{it}}/y_{it}!\Big]) + \log(\theta_2^{\theta_1}\alpha_p^{\theta_1-1}\exp(-\alpha_p\theta_2)/\Gamma(\theta_2))\Big). \tag{41}
$$

A.6 Derivation of the negative binomial (co)variance estimator

Both a derivation of the gradient and of the hessian -in general form given in Appendix A.2- could be used as (co)variance estimate. Here, both forms will be derived. Therefore, the following log likelihood is used:

$$
\sum_i \sum_t \Big[\log\Big(\frac{\Gamma(o+y_{it})}{\Gamma(o)\Gamma(y_{it}+1)}\Big) + y_{it}\log\Big(\frac{\mu_{it}}{\mu_{it}+o}\Big) + o\,\log\Big(\frac{o}{o+\mu_{it}}\Big)\Big]. \tag{42}
$$

The derivations are comparable to the ones given by Ismail and Jemain (2007) (without panel data). Since a derivative to the $\boldsymbol{\beta}$ parameters is taken, it becomes clear that the first log part is removed immediately. The Gradient is as follows:

$$
\begin{aligned}
&\frac{\delta}{\delta\boldsymbol{\beta}} \sum_i \sum_t \Big[\log\Big(\frac{\Gamma(o+y_{it})}{\Gamma(o)\Gamma(y_{it}+1)}\Big) + y_{it}\log\Big(\frac{\mu_{it}}{\mu_{it}+o}\Big) + o\,\log\Big(\frac{o}{o+\mu_{it}}\Big)\Big] \\
&\qquad = \sum_i \sum_t \Big(y_{it}(\frac{1}{\mu_{it}} - \frac{1}{\mu_{it}+o})\boldsymbol{x}_{it} - o(\frac{1}{o+\mu_{it}})\boldsymbol{x}_{it}\Big).
\end{aligned} \tag{43}
$$

The Hessian is as follows:

$$
\begin{aligned}
&\frac{\delta\delta}{\delta\boldsymbol{\beta}\delta\boldsymbol{\beta}'} \sum_i \sum_t \Big[\log\Big(\frac{\Gamma(o+y_{it})}{\Gamma(o)\Gamma(y_{it}+1)}\Big) + y_{it}\log\Big(\frac{\mu_{it}}{\mu_{it}+o}\Big) + o\,\log\Big(\frac{o}{o+\mu_{it}}\Big)\Big] \\
&\qquad = \frac{\delta}{\delta\boldsymbol{\beta}'} \sum_i \sum_t \Big(y_{it}(\frac{1}{\mu_{it}} - \frac{1}{\mu_{it}+o})\boldsymbol{x}_{it} - o(\frac{1}{o+\mu_{it}})\boldsymbol{x}_{it}\Big) \\
&\qquad = \sum_i \sum_t \Big(y_{it}(\frac{-1}{\mu_{it}^2} + \frac{1}{(\mu_{it}+o)^2})\boldsymbol{x}_{it}\boldsymbol{x}'_{it} + o(\frac{1}{(o+\mu_{it})^2}\boldsymbol{x}_{it}\boldsymbol{x}'_{it})\Big).
\end{aligned} \tag{44}
$$

The estimator $\boldsymbol{A}$ is simply calculated by the Hessian and the estimator $\boldsymbol{B}$ as follows:

$$\boldsymbol{B} = \sum_i \sum_t \Big(y_{it}(\frac{1}{\mu_{it}} - \frac{1}{\mu_{it}+o})\boldsymbol{x}_{it} - o(\frac{1}{o+\mu_{it}})\boldsymbol{x}_{it} \Big) \Big(\sum_i \sum_t \Big(y_{it}(\frac{1}{\mu_{it}} - \frac{1}{\mu_{it}+o})\boldsymbol{x}_{it} - o(\frac{1}{o+\mu_{it}})\boldsymbol{x}_{it} \Big) \Big)'. \tag{45}$$

A.7 Derivation of the elasticities of the market share model

The following result will be derived:

$$\frac{\delta M_{it}}{\delta x_{k,jt}} \frac{x_{k,jt}}{M_{it}} = (\delta_{i=j} - M_{jt})\beta_k, \tag{46}$$

where $\delta_{i=j}$ is 1 if i equals j and is 0 if not. Here, the case where i equals j will be derived. First, the derivative will be calculated. Given that $A_{it} = \exp(\mu_i + \epsilon_{it}) \prod_{k=1}^{K} x_{k,it}^{\beta_{k,i}}$ and $M_{it} = \frac{A_{it}}{\sum_{j=1}^{I} A_{jt}}$, it follows from the quotient rule that:

$$\frac{\delta M_{it}}{\delta x_{k,it}} = \frac{(\sum_{j=1}^{I} A_{jt} - A_{it})\exp(\mu_i + \epsilon_{it}) \prod_{g \neq k} x_{g,it}^{\beta_{g,i}} (\beta_{k,i} x_{k,it}^{(\beta_{k,i}-1)})}{(\sum_{j=1}^{I} A_{jt})^2}. \tag{47}$$

Next, the derivative part will be multiplied with the remainder part. This leads to:

$$\begin{aligned} \frac{\delta M_{it}}{\delta x_{k,it}} &= \frac{(\sum_{j=1}^{I} A_{jt} - A_{it})\exp(\mu_i + \epsilon_{it}) \prod_{g \neq k} x_{g,it}^{\beta_{g,i}} (\beta_{k,i} x_{k,it}^{(\beta_{k,i}-1)})}{(\sum_{j=1}^{I} A_{jt})^2} \frac{x_{k,it} \sum_{j=1}^{I} A_{jt}}{A_{it}} \\ &= \frac{(\sum_{j=1}^{I} A_{jt} - A_{it})\exp(\mu_i + \epsilon_{it}) \prod_{g=1}^{K} x_{g,it}^{\beta_{g,i}} \beta_{k,i}}{\sum_{j=1}^{I} A_{jt} A_{it}} = \frac{(\sum_{j=1}^{I} A_{jt} - A_{it}) A_{it} \beta_{k,i}}{\sum_{j=1}^{I} A_{jt} A_{it}} \\ &= \beta_{k,i} - \frac{A_{it}}{\sum_{j=1}^{I} A_{jt}} \beta_{k,i} = (1 - M_{it})\beta_{k,i}. \end{aligned} \tag{48}$$

If i does not equal j, the numerator of the derivative would be $-A_{it}\exp(\mu_i + \epsilon_{it}) \prod_{g \neq k} x_{g,it}^{\beta_{g,i}} (\beta_{k,i} x_{k,it}^{(\beta_{k,i}-1)})$. The denominator would stay the same. Obviously, the result would have been $-M_{it}\beta_{k,i}$.

A.8 Expected value of the random effect

For both the in-sample and out-of-sample forecasts (necessary in the loss functions) of the random effects model we need a value for the random effect. A random effect drawn completely random would not make sense. Also the expected value of the random effect would not make sense. Therefore, we introduce the expected value of the random effect given the data and the parameters: $E[\alpha_i | y_{i1} \dots y_{iT},]$. We get the following:

$$\begin{aligned} E[\alpha_i | y_{i1} \dots y_{iT}, \mathbf{x}_i; \theta] = \int \beta_i g(\beta_i | y_{i1} \dots y_{iT}) = \\ \frac{\int \alpha_i f(y_{i1} \dots y_{iT} | \mathbf{x}_i, \alpha_i; \theta) g(\alpha_i; \theta) \delta \alpha_i}{\int f(y_{i1} \dots y_{iT} | \mathbf{x}_i, \alpha_i; \theta) g(\alpha_i; \theta) \delta \alpha_i}. \end{aligned} \tag{49}$$

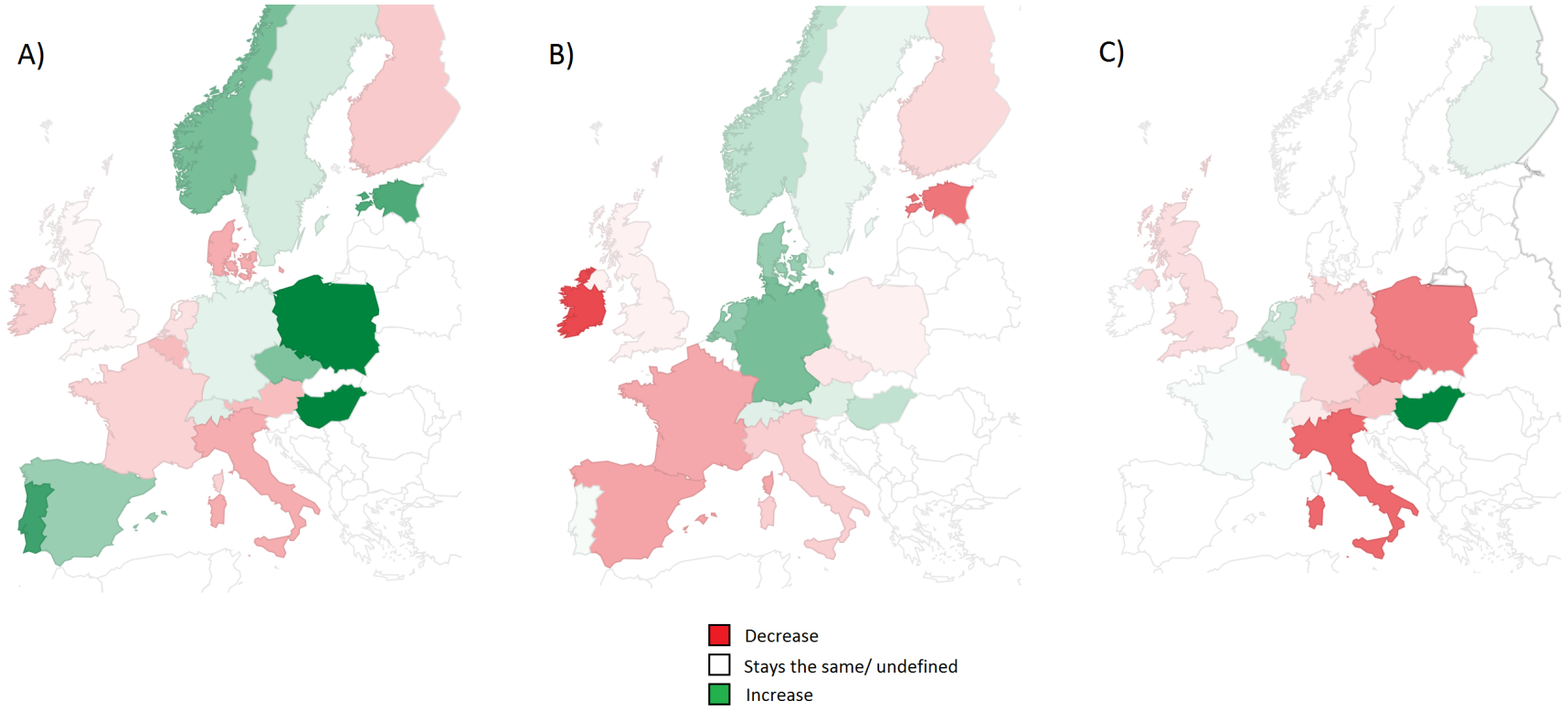
Here the function $f()$ corresponds to the underlying distribution of the model (normal, Poisson and negative binomial). The function $g()$ corresponds to the distribution of the fixed effects (normal, gamma and beta). The integrals are approximated numerically:

$$\frac{\frac{1}{rep} \sum_{rep} \alpha_{rep} f(y_{i1} \dots y_{iT} | \mathbf{x}_i, \alpha_{rep}; \theta) g(\alpha_{rep}; \theta)}{\frac{1}{rep} \sum_{rep} f(y_{i1} \dots y_{iT} | \mathbf{x}_i, \alpha_{rep}; \theta) g(\alpha_{rep}; \theta)}, \tag{50}$$

where rep refers to the number of different α's are drawn from its distribution $g()$.

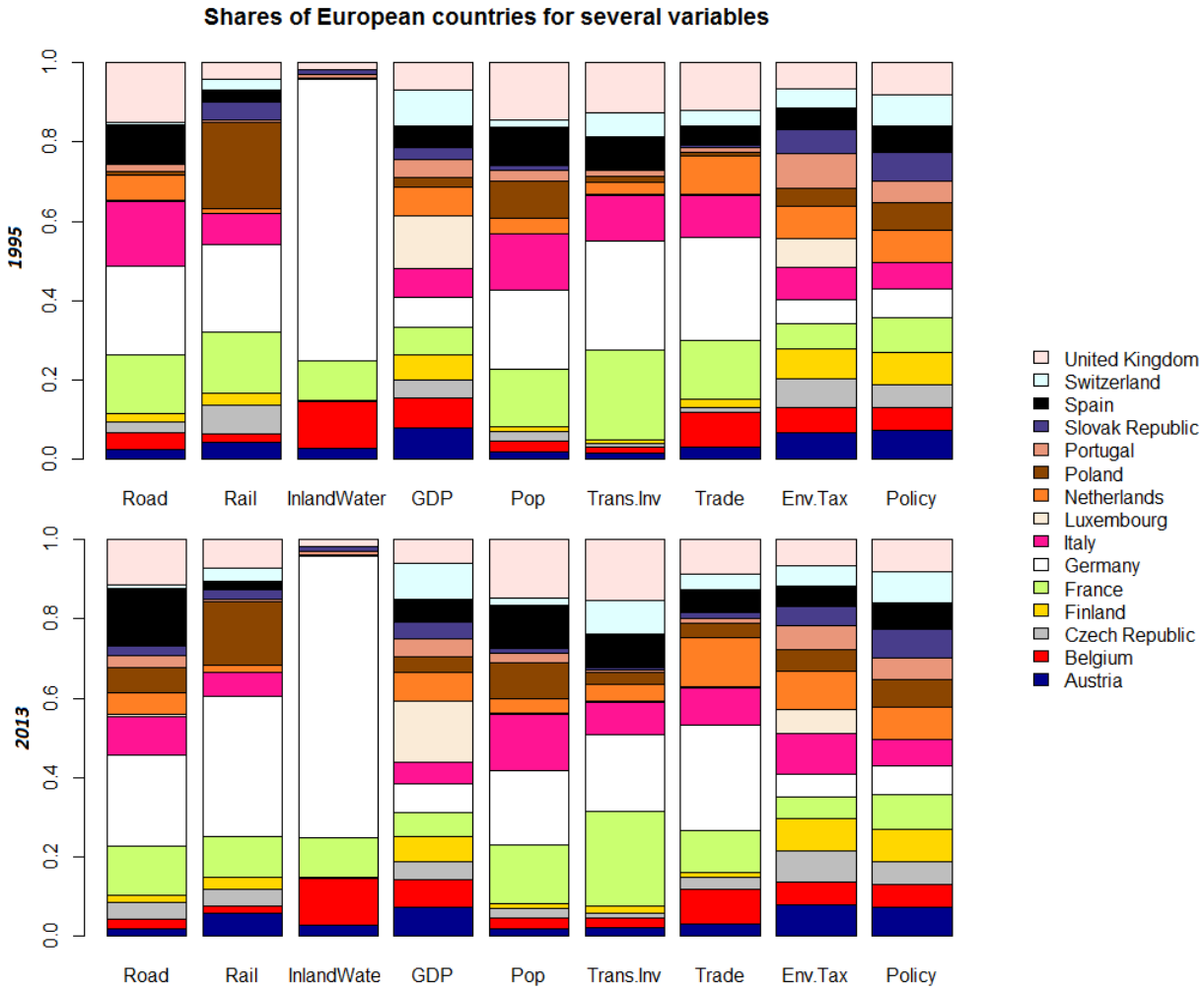
A.9 Change of freight transport

Figure 8: The change of freight transport from 2000 to 2015 for A) Rail transport, B) Road transport and C) Intermodal transport, according to the *Eurostat* data.



A.10 Country shares

Figure 9: The shares of the countries for the variables in the market share model, given for 1995 and 2013.



A.11 Fixed and Random effects

Figure 10: Fixed and random effects are given for the regression on time, weekend percentage, inland water and rail. The fixed effects are plotted in distribution format. The random effects are plotted according to its distribution parameter values.