Estimating Travellers' Preferences for Competition in Commercial Passenger Rail Transport

Johannes Paha, Dirk Rompf, Christiane Warnecke[1]

Abstract

The current level of competition in European commercial passenger rail markets is low and empirical data on customer preferences in intramodal competition has hardly been available, yet. Our study raises the knowledge of competition in commercial passenger rail by exploring the determinants of customers' choice behaviour on two cross-border routes, Cologne-Brussels and Cologne-Amsterdam. We analyse stated preference information from about 700 on-train interviews by means of multinomial Logit regressions. Our analysis indicates that customers experiencing competition (Cologne-Brussels) show a higher preference for competitive services than customers for whom competition is a purely hypothetical situation (Cologne-Amsterdam). Moreover, travellers show a status quo bias, i.e. a preference for the service provider on whose trains they were interviewed which partly stems from switching costs. These findings regarding status quo bias and switching costs complement previous studies on the outcome of intramodal competition, implying that entry is even more difficult than they predicted.

Keywords: Competition, Passenger, Rail, Transport, Discrete Choice, Multinomial Logit

JEL Codes: C25, D12, D40, L92

[1] We gratefully acknowledge the cooperation of Deutsche Bahn, Thalys, SNCF and SNCB who permitted and supported the on-train surveys. We would like to thank the students of International School of Management (Frankfurt/Main) for carrying out the passenger-interviews.
Valuable comments on this paper were provided by Georg Götz, our colleagues at Justus-Liebig-University Giessen and Chris Nash. We would also like to thank the discussants and audiences at Workshops in Münster and Giessen, Khumo Nectar Conference 2011 and EARIE 2011.

1 Introduction

Since the 90ies commercial passenger rail markets in Europe have been opened to competition under open access regimes within several countries, e.g. in Germany, UK, Austria, Italy, Poland and Sweden. In 2010 international services in Europe have also been liberalised. In spite of open markets, very little entry and almost no on-track competition has been seen, yet. This could change soon: entrants are preparing for hourly services with new trains parallel to incumbent operators in Italy (NTV) and Austria (WESTbahn) starting in 2011/2012.

The outcome of open access competition in passenger rail markets is uncertain. Since the 1990ies several researchers have studied the effects of competition and entry in this sector (for example, see Preston et al. (1999) and Preston (2009), Steer Davis Gleave (2004), Ivaldi and Vibes (2008), Friederiszick et al. (2009)). As a general outcome, the theoretical and empirical studies suggest that intramodal competition is possible on a low scale. The characteristics of the market, i.e. strong economies of density, potential economies of scale, network effects and vehicle investments with a high risk of sunk costs make entry difficult. The studies diverge regarding their conclusions on the attractiveness of commercial passenger rail markets for entrant rail companies. They suggest that entrants might concentrate on cream-skimming (Preston et al. 1999), low-cost services (Ivaldi and Vibes 2005), intercity services (Friederiszick et al. 2009) or international passenger rail services (Steer Davis Gleave 2004).

We contribute to this literature by, first, collecting a new set of stated customer-preference data regarding intramodal competition on two cross-border routes (Cologne-Brussels and Cologne-Amsterdam). Our study is based on almost 700 interviews where we collect socio-demographic data of the interviewees as well as information concerning their trip at the time of the interview. The main part of the questionnaire consists of 21 stated preference scenarios in which travellers are asked to bring the services of two competing rail companies (incumbent and entrant) and a cooperative offer into a rank-order. These scenarios differ in the absolute and relative prices of the three alternatives. We choose Cologne-Brussels as the only route in Europe with significant and established on-track competition and Cologne-Amsterdam as reference case with a cooperative service.

Second, we econometrically analyse and compare the preferences of travellers by estimating a multinomial logit discrete choice model for competitive and cooperative markets as has become standard in transport modelling (see e.g. Louviere et al. 2000). From these regressions we determine the respondents' mode choice probabilities, i.e. choosing the service of the entrant, the incumbent, or a cooperation of the two. These results are in line with expected behaviour based on existing demand studies as outlined e.g. by Oum (1990). First class and business passengers are, for example, found to have an above average preference for frequent connections as one indicator for comfort.

We identify price as a major influencing factor for customer choices whose effects vary across different groups of customers. Again business and first class passengers are found to be less price sensitive than leisure and second class travellers.

Third, the results of these regressions allow us to identify customers' costs of switching between the rail service operators presented in the stated preference scenarios. We formulate a set of hypotheses to assess the relevance of specific switching costs on customer preferences (see e.g. Klemperer 1995). We consider the effect of loyalty programmes, frequency of rail travel, actual trip with/without transfer and whether a traveller purchases his/her tickets over the internet. Most of these hypothesised effects can be substantiated with the strongest effects coming from the possession of customer cards. Our analysis also suggests that interviewees who experience competition (Cologne-Brussels) show a different answering behaviour than travellers for whom competition is a purely hypothetical situation (Cologne-Amsterdam).

We also find a marked tendency on both routes that respondents choose the service they are using at the time of the interview. In terms of switching costs this could be due to costs of learning to use a new service, psychological switching costs or transaction costs. Samuelson and Zeckhauser (1988) explain that a "status quo bias" of customers may result from both rational reasoning and psychological effects. Habitual travel choice provides a further explanation for the preference of the known service provider (Gärling and Axhausen 2003).

In line with these findings we show for the route Cologne-Amsterdam that interviewees are relatively reluctant towards choosing the hypothetical entrant RailX. Interviewees have never heard of this presented entrant before and are not given additional information on the company. In comparison to the entrant on the Cologne-Brussels route (i.e. the well known Deutsche Bahn AG) we identify a bias in favour of the better known company (see for example Hensher 2009). This may have implications for predictions about the profitability of entry by an unknown operator. Stated preference analyses as ours might be useful for predicting the market share of an unknown entrant shortly after its entry onto a route. However, they might not be indicative of long-term market shares.

These points are illustrated in greater detail below. Section 2 provides our theoretic analysis of the competitive effects in commercial long distance passenger rail transport and of the approach to demand reactions and customer preferences. The survey design and collected data is presented in section 3. In section 4, we provide a utility function for the current problem of choosing the utility-maximising train service and present the regression-results from estimating this utility function with the collected data. Based on the regression results in section 5 we calculate choice probabilities for cooperation and entrant and discuss the effects of switching costs. Section 0 concludes.

2 Related Literature

In this section, we show how our paper relates to existing literature on commercial passenger rail transport. Moreover, we give an overview on relevant (psychologically motivated) effects that arise when customers choose among competing products.

2.1 Competition in Commercial Passenger Rail Transport

In this section we present the theoretical foundations of customer choices and of competition in long distance passenger rail transport. Studies on the effects of competition in commercial passenger rail concentrate on the analysis of intramodal competition or look at rail as one possible choice in intermodal competition. (see, e.g., Ivaldi and Vibes 2008). In the latter models, a reduction in the prices of two competing rail companies does not only affect these companies' relative market share (intramodal effect). The reduction in prices is also likely to attract customers who, otherwise, would have taken the plane or a car (intermodal effect). These types of analysis are nested in each other insofar as a traveller can be modelled to, first, decide on taking the train or not and, second, choosing a rail-operator (see, e.g., Preston et al. 1999: 83).

A comprehensive study on passenger rail competition was carried out by Preston et al. (1999) who perform a simulation-analysis on the potential for on-track competition in the UK passenger rail industry and its welfare effects. They focus on the intramodal effects, but take intermodal shifts into account. On the demand side they use stated and revealed preference data obtained from a survey looking at customer choices of preferred departure times, ticket types, class and mode of travel on a specific UK route. The analysed scenarios vary in the price levels of entrant and incumbent and in the interavailability of tickets. They find that a few entry scenarios could be attractive, e.g. cream-skimming and niche entry through product differentiation, but also conclude "that on-track competition can increase benefits to users but usually reduces welfare because of greater reductions in producer surpluses" (Preston et al. 1999: 92).

The same model was used for simulation of competition in the Swedish market (see Preston 2009), with a major difference being lower track access charges. For Sweden the model indicates that head-on competition could also be commercially feasible on the important intercity routes, although the welfare maximising scenario would be a regulated monopolist incumbent with fare reductions and frequency increase.

Steer Davis Gleave (2004) use a similar approach for modelling the effects of intramodal competition on two international high-speed routes and for assessing the overall impact of competition in European high speed rail. Based on the model results, on case studies and stakeholder views they see potential for competition in international passenger rail services with cabotage.

Ivaldi and Vibes (2008) use a game theoretic simulation model for analysing intramodal *and* intermodal (rail, road, and air) competition. On the demand side, this model is based on discrete choice theory. In section 4 below, we present a similar, non-nested (i.e. for intramodal competition only) demand model. On the supply-side, firms are assumed to compete in prices á la Betrand. Due to the lack of sufficient data, Ivaldi and Vibes (2008) do not *estimate* their model but instead *calibrate* it to the link Cologne-Berlin in Germany. Based on this calibrated model they evaluate the effect of structural and regulatory changes. For intramodal competition they conclude that low-cost entry could be viable and beneficial to consumers.

Friederiszick et al. (2009), on the one hand, look at intermodal competition between air and rail travel. On the other hand, they analyse intramodal competition in a separate approach. In the latter analysis they explore the entry potential on 207 national and international routes starting or ending in Germany. Based on a profitability analysis for the routes they conclude that entry could be attractive on intercity routes. However they do not consider strategic interactions between entrant and incumbent and do not take intermodal shifts into account.

The demand models of the discussed studies are usually based on a few influencing factors like price, travel time, departure time and specific quality indicators. These factors have a major influence on customer choice in an inter- and intramodal perspective. However, the choice between intramodal operators could also be influenced by switching costs and some psychological persistence effects which have been documented for other industries but have not been researched for the passenger rail market yet. We focus our analysis on the impacts of these factors. With our empirical data we thus provide a novel view on the effects of competition in commercial passenger rail transport. The effects which we identify should be seen as complementing previous studies. They could affect the market potential of firms as well as welfare effects of competition.

2.2 Customer Choice Behaviour for Competing Products

The impact of switching costs on competition has largely been discussed for other industries, for example, by Farrell and Klemperer (2007) and by Klemperer (1995). Seabright et al. (2003: 60) point out for the passenger rail sector, that “switching costs (natural or strategic) should be a strong limitation to competition”. Switching costs are defined by Klemperer (1995) as costs that a consumer incurs from switching to a competitor’s product even when the products of the two firms are functionally equal. He identifies several types of costs potentially involved with switching service providers, i.e. need for compatibility with existing equipment/services, transaction costs of switching suppliers, costs of learning to use new brands, uncertainty about the quality of untested brands, effects of discount coupons and similar devices and non-economic brand-loyalty. In the passenger transport sector, switching costs have mainly been analysed for airline competition. Studies focus strongly on the impact of customer loyalty programmes and brand-loyalty (see for example Banerjee and Summers 1987, Carlsson and Löfgren 2004) and identify substantial switching costs.

In this context, habitual behaviour, i.e. the repeated performance of behaviour sequences (Gärling and Axhausen 2003), is a further factor that needs to be considered. Habitual behaviour may be a strong driver of travel choices. It could be seen as deliberate or as non-deliberate repeated actions. In the first case customers would decide rationally, in the second case they would not think about their decision but make script-based choices. Habit as deliberate action could partly be based on switching costs as for example transaction and learning costs or uncertainty with regard to untested brands.

Behavioural economics can give further explanations for customer preferences under competition. Samuelson and Zeckhauser (1988) propose that consumers are disproportionately likely to stick with the status quo. They explain the "status quo bias" with rational decision making in the presence of transaction cost and/or uncertainty (note: link to switching costs), with cognitive misperceptions and psychological commitment. Samuelson and Zeckhauser (1988: 10) suggest that the "individual may retain the status quo out of convenience, habit or inertia, policy (company or government) or custom, because of fear or innate conservatism, or through simple rationalization." They also argue that in comparison to laboratory experiments this effect should be much stronger for decision makers in the real world, for example because the experiment presents alternatives to customers they might not have thought about in the first place.

The status quo effect thus describes the general effect of preferring the actual/known product. It includes the above described effects of switching costs and habitual behaviour. Such an effect is well known to stated preference analysis in the transport sector. It is often discussed as "inertia" effect and strongly explained by habitual behaviour (see e.g. Goodwin 1977 for a discussion of habitual choice and Train 2009).

In our stated preference experiment we expect that a status quo bias would lead to an especially high preference for the rail service operator on whose trains the travellers were interviewed (see also Cantillo et al. 2007: 196). They also point out that the same behaviour could prevail after a change in circumstances, as a change might involve costs, both objective and psychological. Thus, we would find a status quo bias also on the route with established competition.

Summing up, as a contribution to the analysis of open access competition in passenger rail, our study looks at customer choices in intramodal competition with a strong focus on the perception of switching costs and on other psychological influences on travel choice as explained above which could lead to a status quo bias. Based on the comparison of two routes – one currently with competition, the other with a cooperative service – we aim to identify patterns of rail customers' behaviour in order to predict their likely reaction to competition.

3 The Data

3.1 Study Design and Data

In this section, we provide information on the study design and the data collected. The data set consists of stated and revealed preference data as well as socio-demographic information of almost 700 passengers on trains of Deutsche Bahn and Thalys on the two international routes Cologne-Brussels and Cologne-Amsterdam. The passenger-interviews were conducted by students of the International School of Management in Frankfurt/Main during two weeks in May 2010.

The routes Cologne-Brussels and Cologne-Amsterdam were chosen because they share the characteristics of being international routes with a similar distance and a similar mix of customers, e.g. business and leisure travellers. The main difference of the two routes is the fact that rail customers on the Brussels-route already experience competition between Thalys and Deutsche Bahn. In contrast, on the Amsterdam-route passenger rail services are provided by Deutsche Bahn and Dutch Railways (NS) in cooperation. This allows us to study the effect of experience with competition on the preferences of the interviewees.

At the time of the interview, competition between Thalys and Deutsche Bahn takes the following form: Customers face two generally separate information and sales systems (only DB systems inform about Thalys trains and partly sell tickets, but not vice versa), independent pricing systems and different customer retention programs. In case of delay, customers cannot easily access a train of the other company. Coordination only occurs in some areas such as schedule planning.

The market survey consisted of two parts:

(1) We collected data on the socio-demographic characteristics of the customers as well as characteristics of their journey and ticket at the time of the interview.
(2) We collected stated preference data on the preferences of the interviewees in 21 scenarios of intramodal competition between an entrant and an incumbent.

The socio-demographic and general travel characteristics of the customers include gender, age, the frequency of travels on this route during the last 12 months and the possession of a discount/loyalty card. Other characteristics are related to the specific journey of the decision maker. These are the purpose of the journey, the chosen class and the distribution channel. An overview on the definitions of these variables is provided in the Appendix.

In the second part of the interview, the rail passengers were asked to rank – i.e. to make a discrete choice – three types of train-service on the respective route for different combinations of prices (see Figure 1 for an example). The *J* = 3 alternatives in the choice set of traveller *n* are:

(1) Service provided by an entrant in a situation with competition
(2) Service provided by an incumbent in a situation with competition
(3) Integrated service provided in cooperation by the two railways

This choice set satisfies the criteria of choice sets for discrete choice analysis (Train 2009: 11). First, the alternatives are *mutually exclusive* because choosing one alternative implies not choosing any of the other alternatives. Second, the choice set is *exhaustive* in that all possible alternatives of making a rail-travel are included. This implies that the below results on customers' preferences must be interpreted conditional on their previous choice of taking the train instead of other modes of transportation. Third, the number of alternatives is *finite*, which is the defining characteristic of a discrete choice model.

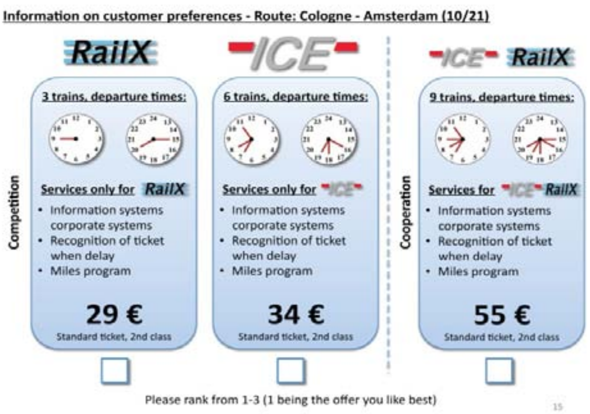


Figure 1: Example of stated choice scenarios

Interviewees on both routes were given the information that the difference between cooperative and competitive services lies in integrated or separate (i) information and sales systems, (ii) ticket acceptance in case of train disturbances, and (iii) bonus programs. On the Brussels-route, Thalys is the incumbent and DB the entrant. On the Amsterdam-route, DB is presented as the incumbent with a hypothetical firm RailX being the entrant. Hence, competition on the route Cologne-Amsterdam constitutes a hypothetical situation. On both routes the incumbent is assumed to provide a higher number of services than the entrant (6 vs. 3 train pairs per day). This depicts the situation with Thalys and Deutsche Bahn on the Brussels- Cologne itinerary for the train schedule in 2010. The number of integrated services by the cooperation equals the sum of services of entrant and incumbent, i.e. there is no change in the total number of services between cooperation and competition.

The alternative "cooperation" in the SP experiment represents the dominant way international services are currently organised in Europe. The involved railways usually cooperate, providing a joint monopoly service. For the competitive alternatives we assume a specific situation with infrastructure bottlenecks impeding an increase in frequency. Entry would only be possible if the entrant replaces services of the incumbent. This is a probable situation on many national and international connections in Europe due to infrastructure bottlenecks on links or in nodes. We also use the fixed total number of services as a simplification to identify how customers value the presented potential drawbacks from competition (e.g. disintegration of services on the marketing and sales side) compared to cooperation. In many cases on-track competition would lead to benefits with regard to higher frequency and other service improvements which we do not consider here. Frequency increase has for example been included in the simulations by Preston (Preston et al. 1999, Preston 2009).

In the stated preference scenarios, the price of the incumbent (p_i) and that of the entrant (p_e) varies in the interval between the regular price for a second class-trip on each route and the lowest price for special offers. The cooperative price (p_c) is kept constant at the value of the second class regular price charged at the time of the survey. The incumbent never charges a price below that of the entrant. In order to cover a large number of possible price variations, 21 scenarios were defined. The below table shows that on the Brussels route prices are varied in six steps of either 5€ or 6€. For the Amsterdam route prices (in parentheses) are varied in six steps of up to 10€.

p_i — p_e	45 (55)	39 (45)	34 (39)	29 (34)	24 (29)	19 (19)
45 (55)	scen. 1	scen. 7	scen. 12	scen. 16	scen. 19	scen. 21
39 (45)		scen. 2	scen. 8	scen. 13	scen. 17	scen. 20
34 (39)			scen. 3	scen. 9	scen. 14	scen. 18
29 (34)				scen. 4	scen. 10	scen. 15
24 (29)					scen. 5	scen. 11
19 (19)						scen. 6

Table 1: Price scenarios for Cologne-Brussels and Cologne-Amsterdam (in parentheses)

3.2 Data Quality

How well is real market behaviour captured by stated preference data? Bates (1998) and Wildert (1998) discuss some issues with regard to data quality of stated preference experiments that can be summarised to the following two questions: How well does the scenario setting relate to the interviewees' actual decisions (design of the study)? How was the attitude towards answering the questionnaire (inclination to respond)? We analyse these two questions in turn.

Regarding the design of the questionnaires used in our study, we chose a form that closely resembles the concrete situation that customers face when making their travel decisions. After deciding for travelling by rail, customers would search for train departure times and compare the prices of the (hypothetical) rail companies operating on the desired route either competitively or cooperatively as is reflected in our questionnaires. We deliberately did not choose a two level approach in which the customers would, first, have been asked if they prefer competition or cooperation and, second, would have had to decide between entrant and incumbent in a competitive situation. We decided against this approach as the decision between cooperation and competition is quite abstract and, thus, might have lowered the quality of answers and increased a hypothetical bias.

Concerning interviewees' inclination to respond, including 21 scenarios must not be considered too extensive for our stated preference analysis. Louviere and Hensher (2001: 129) point out: "There is little agreement on what "complicated" or "lengthy" means and little empirical evidence to support this conventional wisdom". They list several studies that did not find significant effects from "lengthy" stated preference experiments on response means or model parameters, but rather on differences in random component variance. We decided to use a high number of scenarios because the complexity of our study is low, as only one attribute (price) is changing. Additionally, pre-tests showed that the questionnaires could be answered in reasonable time. In total it took respondents 15 to 20 minutes to answer the questionnaire.

With idle time during their train journey, the travellers were positively inclined to answering the questions. Only in very few cases, we observe effects like weariness that might have affected their answers. As such effects are unsystematic, they may be expected to only increase the variance of the error term in our below regressions. As a result, economically plausible effects might not be found to be statistically significant. Since the number of observations in our regressions is large and the number of possibly impaired observations is so small, we do not find any indication of an effect on statistical significance.

4 Empirical Analysis

In the following, we present the design and the results of our empirical analysis of the above data. In particular, we analyse the process of decision making of the interviewed passengers by performing multinomial logit regressions for discrete choices. These are based on the assumption that the interviewed passengers ranked the three alternatives – i.e. *rival*, *Deutsche Bahn*, or *cooperation* – according to their utility function. The alternative that yields the interviewee the highest utility receives rank 1 and is assumed to be the respondent's observed choice in a real-world situation. In section 4.1, we present the utility function and the underlying hypotheses for our specific decision problem. Section 4.2 presents the regression results for this utility function and draws economic inferences from this regression.[2]

4.1 The Utility-Function and its Underlying Hypotheses

We assume that decision maker *n* obtains utility U_{nj} from alternative *j*. *j*=1 applies if the decision maker chooses the *rival* of Deutsche Bahn, i.e. Thalys on the Brussels-route and RailX on the Amsterdam-route. *j*=2 applies for choosing the service of *Deutsche Bahn*, and *j*=3 applies for the *cooperation* offer. On basis of this *unobservable* utility function an interviewed passenger makes his/her *observable* ranking-decision and chooses alternative *j* if and only if $U_{nj}>U_{ni}\ \forall\ j\neq i$.

The basic idea of discrete choice modelling is to infer information about the interviewees' unobservable decision making process by analysing their observable choices. This is done by

[2] A more technical description of discrete choice models and the regression tables is presented in Appendix 8

specifying utility U_{nj} as a function that relates the three above alternatives to observable attributes of both the alternatives and of the decision maker. The economic reasoning for including certain attributes, labelled x_{nj} $\forall$ j, is explained further below in this section. They enter utility function (1) in form of the column-vector $\boldsymbol{x}_{nj}$.

$$U_{nj} = \boldsymbol{\beta}_j{}' \boldsymbol{x}_{nj} + \varepsilon_{nj} \tag{1}$$

Given the assumption that the error term ε_{nj} is independently, identically distributed extreme value one can calculate the probability of a particular decision maker to prefer either of the alternatives j to the others. The column-vector of coefficients $\boldsymbol{\beta}_j$ is estimated as the set of parameters that maximises the likelihood to observe the decision makers' choices as stated in the interviews. We assume a utility function that applies to all customer groups. Therefore, $\boldsymbol{\beta}_j$ is not indexed by *n.* Note that econometrically estimating this utility function does not mean to cardinally measure utility, which would be at odds with utility theory. Rather, the estimated function is a means to predict the probability that a certain type of customer would choose one of the alternatives j. The specification of the utility-function in non-matrix form looks as follows.

$$\begin{aligned}
U_{nj} = \; & \beta_{j,1} + \beta_{j,2} \cdot D_{\text{class=1}} + \beta_{j,3} \cdot D_{\text{business}} + \beta_{j,4} \cdot D_{25} \\
& + \beta_{j,5} \cdot D_{6+} + \beta_{j,6} \cdot D_{\text{DB}} + \beta_{j,8} \cdot D_{\text{discount_r}} \\
& + \beta_{j,21} \cdot D_{\text{direct}} \\
& + \left(p_r / p_c \right) \cdot \Big[\beta_{j,9} + \beta_{j,10} \cdot D_{\text{discount_r}} + \beta_{j,12} \cdot D_{\text{business}} \\
& \qquad + \beta_{j,13} \cdot D_{\text{internet}} + \beta_{j,22} \cdot D_{\text{direct}} \Big] \\
& + \left(p_{DB} / p_c \right) \cdot \Big[\beta_{j,14} + \beta_{j,16} \cdot D_{DB25} + \beta_{j,17} \cdot D_{DB50} \\
& \qquad + \beta_{j,18} \cdot D_{DB100} + \beta_{j,19} \cdot D_{\text{business}} + \beta_{j,20} \cdot D_{\text{internet}} \\
& \qquad + \beta_{j,23} \cdot D_{\text{direct}} \Big] \\
& + \varepsilon_{nj}
\end{aligned} \tag{2}$$

We start with including a constant associated with coefficient $\beta_{j,1}$. This constant ensures that the unobserved part of utility ε_{nj} has zero mean by construction. $\beta_{j,1}$ measures the overall likeliness of a group of passengers to prefer the *rival* ($\beta_{1,1}$>0) or *Deutsche Bahn* ($\beta_{2,1}$>0) to the cooperative offer. Moreover, we include a set of dummy variables to control for the possibility that different groups of passengers differ in their non-price related inclination to choose either of the alternatives. Dummies[3] are included for the following groups of customers, which are sometimes overlapping:

[3] As is standard, a dummy variable takes a value of 1 if interviewee *j* belongs to a certain group of customers and 0 otherwise. An overview on the definition of variables is provided in the Appendix.

(1) First-class passengers ($D_{class=1}$)

(2) Business travellers ($D_{business}$)

(3) Frequent travellers (D_{25} and D_{6+}, i.e. an interviewee has travelled 2-5 times or at least 6 times on this route over the 12 months prior to the interview)

(4) Holders of a discount card of Deutsche Bahn (D_{DB}) or its rival ($D_{discount\_r}$)

(5) Travellers who (at the time of the interview) only travel between places on the routes Cologne-Amsterdam or Cologne-Brussels without the need to switch trains (D_{direct})

Concerning the signs of the coefficients β of these variables we have the following expectations. These are formulated as crosschecks or hypotheses. The earlier should be satisfied in order to show the general plausibility of our calculations. The latter are at the focus of our research. Accepting or rejecting these hypotheses helps to better understand the demand characteristics of commercial passenger rail transport.

Hypothesis 1: We expect the coefficient of the dummy variable for first-class passengers ($D_{class=1}$) to be negative ($\beta_{j,2}<0$). This is because first-class passengers are expected to have a high preference for comfort as has, e.g., been shown by Oum et al. (1986) for the demand for air travel. One aspect of comfort is the convenience to have a high number of connections on a route. The number of connections per day is highest for the cooperative offer.

Hypothesis 2: Business travellers ($D_{business}$) are expected to have an above-average preference for frequent connections ($\beta_{j,3}<0$). This is because it is more important for business travellers to keep their appointments than for leisure travellers. Whitaker et al. (2005) show for the airline sector that business travellers are much less flexible on departure/arrival times and value a preferred schedule four to five times higher than leisure travellers.

Hypothesis 3: Customers who possess a customer card are (at least to some extent) locked in to the issuer of the card. BahnCard-holders (D_{DB}) are expected to have a reduced preference for the *rival* of *Deutsche Bahn* ($\beta_{1,6}<0$). Customers who hold a discount-card of the *rival* ($D_{discount\_r}$) are expected to have a reduced preference for *Deutsche Bahn* ($\beta_{2,7}<0$).

This expectation of a positive relationship between customer loyalty programmes and the market share of the respective company is in line with, e.g., Bolton et al. (2000), Rust et al. (2000) and Verhoef (2003). They show that customer loyalty programmes have a significant positive effect on customer retention. For the airline sector, empirical studies (e.g. Carlsson and Löfgren 2004) conclude that members of loyalty programmes have higher switching costs.

Hypothesis 4: We expect customers with a discount-card to have a higher preference for the cooperative offer as compared to an isolated offer by the issuer of the card, i.e. *Deutsche Bahn* ($\beta_{2,6}$<0) or its rival ($\beta_{1,7}$<0). This is because in case of cooperation a discount card can be used on the trains of both competitors. Thus, in the cooperative case a discount card would be of greater value to customers.

Hypothesis 5: Passengers who (at the time of the interview) travel on a direct train on the two routes (D_{direct}) are expected to have a higher preference for competition ($\beta_{j,21}$>0) than customers on itineraries beyond Brussels, Cologne and Amsterdam that involve a change of trains. We expect that travellers on direct trains have lower transaction costs and could more easily switch between the offers of the competitors. In contrast, travellers who need to switch trains may have an additional preference for buying a single ticket for their entire trip. Their preference for competition might possibly be lower because of an additional preference for exploiting network effects.[4]

The variables D_{25} and D_{6+} indicate travellers (D_{25}) who have travelled on this route 2-5 times over the 12 months prior to the interview and those (D_{6+}) who have travelled on this route six and more times. We expect different effects for frequent travellers. They could (i) show an especially high habitual behaviour, (ii) reveal high experiences with competitive or cooperative market situations or (iii) have a preference for high service frequencies. In this context Seabright et al. (2003) assume that frequent travellers may have high switching costs, while occasional travellers are unlikely to have switching costs. Because of this multitude of effects we do not have specific expectations concerning the signs of the coefficients $\beta_{j,4}$ and $\beta_{j,5}$ that are associated with these frequent traveller variables but include D_{25} and D_{6+} in order to prevent an omitted variables bias.

In addition to these fixed effects, the utility depends on the price of the *rival* (p_r) and the price of *Deutsche Bahn* (p_{DB}). Therefore, we include these variables in the specification of utility function (1). Note that prices are not included in levels but relative to the cooperative price (p_c), i.e. we include p_{DB}/p_c and p_r/p_c. This scaling does not affect the interpretation of the estimates because the cooperative price is the same in all 21 scenarios. However, the scaling facilitates the comparison of the regression-results for the route Cologne-Amsterdam (with p_c=55) and the route Cologne-

[4] On the route Cologne-Brussels there are two further relevant groups of customers. First, passengers who travel from Cologne to Paris. This route is served as a direct train by Thalys. Second, passengers who travel from Frankfurt to Brussels. This route is served as a direct train by Deutsche Bahn. Our empirical results show that these two groups answered the questionnaire no different than passengers whose journey goes beyond Cologne/Brussels and who, thus, have to switch trains in between. Therefore, we do not further distinguish between customer-groups according to their destinations.

Brussels (with p_c=45). Regarding the coefficients associated with these variables we formulate hypothesis 6.

Hypothesis 6: The own-price elasticity of rail-travel is negative. The utility of choosing either alternative depends negatively on its own price ($\beta_{1,9}<0$ and $\beta_{2,14}<0$). The related price effect of competition may be considered particularly important in book-ahead markets such as long distance rail services (Preston 2009: 4) as analysed in this paper. Empirical studies on price elasticities are for example summarised by Oum et al. (1990).

Prior research as well as pre-tests with our dataset suggest that interviewees' sensitivity towards price changes varies across different groups of customers. Therefore, we control for these effects by including interaction terms where relative prices are multiplied with dummy variables that indicate the respective groups of customers.[5] Regarding the related coefficients we have the following expectations.

Crosscheck 1: We expect business travellers ($D_{business}$) to be less price sensitive than leisure customers because they often do not pay for the trip themselves ($\beta_{1,12}>0$ and $\beta_{2,19}>0$). This expectation is confirmed by a large number of studies (see Oum et al. (1990) for an overview).

Hypothesis 7: We expect that passengers who have bought their ticket on the internet ($D_{internet}$) react more sensitively to changes in the price of a ticket ($\beta_{1,13}<0$ and $\beta_{2,20}<0$). This hypothesis is based on the assumption that internet users are inclined to search for best prices and are used to check with a number of service providers. They can be expected to have less transaction and learning costs and be less cautious with the quality of untested brands.

Crosscheck 2: We expect that travellers who possess a loyalty card are less sensitive to changes in the price of the service of the card's issuer ($\beta_{2,16}>0$, $\beta_{2,17}>0$, $\beta_{2,18}>0$, and $\beta_{1,10}>0$). The loyalty cards included in the study provide benefits/rewards to frequent travellers (Thalys The Card) and discounts (DB BahnCards, NS RailPlus).

[5] For example, the coefficient $\beta_{1,9}$ measures the sensitivity of a customer's decision to choose *rival* when the price of the rival is varied and when the customer is of the following type: He/she

does not possess a discount card of the rival ($D_{discount\_r}$=0),
is a private traveler ($D_{business}$=0),
did not buy the ticket over the internet ($D_{internet}$=0), and
considers the route Cologne-Amsterdam or Cologne-Brussels only part of his/her current journey (D_{direct}=0).

The price sensitivity of an otherwise identical business traveler is measured by the sum of the coefficients $\beta_{1,9}$ and $\beta_{1,12}$.

Crosscheck 3: The price-sensitivity should be the lower the higher the discount of the card ($\beta_{2,16}<\beta_{2,17}<\beta_{2,18}$). The DB BahnCard can be bought with 25%, 50% or 100% discount. We assume that the level of discount influences the price sensitivity of the respondents.

Crosscheck 4: In accordance with Hypothesis 5, we expect that passengers who travel on a direct train between Cologne and Brussels (or Amsterdam) react more sensitively to price-changes than customers who have to switch trains ($\beta_{1,22}<0$ and $\beta_{2,23}<0$).

Although the services by Deutsche Bahn and its rival are expected to be substitutes, which implies positive cross-price elasticities, the coefficients $\beta_{2,9}$ and $\beta_{1,14}$ need not necessarily take positive signs. This is because, e.g., an increase in the price of Deutsche Bahn makes customers switch to the rival and to the alternative cooperation. Finding $\beta_{1,14}>0$ would indicate that a greater portion of passengers switches to the entrant than to the cooperation. We do not have any expectations concerning this switching behaviour.

In section 4.2 we present the results of the estimation of utility function (2) for the two routes Cologne-Amsterdam and Cologne-Brussels. In the latter case we run the regression for the two datasets collected on Thalys-trains and on DB-trains. The regressions support the vast majority of the above hypotheses.

4.2 The Results of the Regression

We estimate the above model by means of multinomial logit-regressions. We only include business and leisure travellers in the regressions, because we expect commuters to have a different utility function in comparison to passengers who only occasionally travel on this route. In our regressions we find some evidence for this effect that is likely to distort our estimates. Therefore, the below results apply to non-commuters only, who account for the vast majority of respondents in our samples (94% on average). The tables of the estimated regressions are provided in Appendix 8.[6]

In this section, we interpret the signs and significance levels of the estimated effects with regard to the above hypotheses and crosschecks. We first interpret the estimated **fixed effects**. First-class passengers are found to prefer cooperation (i.e. frequent connections and integrated marketing services) to either of the competitive offers (hypothesis 1). However, this effect is not always statistically significant. In all cases, business travellers show a significant preference for cooperation (hypothesis 2). Concerning the effect of customer loyalty cards, we find that holding a BahnCard or a Thalys-card significantly affects customers' choices. The stated preference analysis of travellers who

[6] Note that the absolute values of the coefficients in these regressions cannot be compared because the coefficients are scaled by the standard error of the unobserved factors. Coefficients and standard errors are not separately identified and the standard errors differ for each two regressions (Train 2009: 41).

were interviewed on DB-trains supports the propositions that holding a BahnCard (i) reduces the probability to choose the *rival* and (ii) increases the interviewees' preference for *cooperation* ($\beta_{1,6}<\beta_{2,6}<0$; hypotheses 3 and 4). The holders of the Thalys loyalty card show an even stronger preference for cooperation. They represent a small group of interviewed customers (max. 5%) who go by train on the route relatively frequently. Please note that the aggregate effect of holding a customer card cannot be determined solely with regard to these fixed effects but requires taking into account the effect of the discount cards on price-sensitivity. The evidence regarding customers, who have their starting point and destination in or between the two cities Cologne-Amsterdam respectively Cologne-Brussels and, thus, are expected to switch between operators relatively easily, is inconclusive (hypothesis 5). The estimated coefficients for frequent travellers do not show a clear pattern. However, we can neither reject the presumption that frequent travellers show habitual travel choice behaviour.

The **own-price effects** are negative and statistically significant at the 1%-level in all regressions (hypothesis 6). In the case of Brussels the own price effect is considerably higher for the incumbent which could be explained by its higher service frequency. Except for one case,[7] we find that business customers have a lower price-sensitivity than leisure travellers (crosscheck 1). BahnCard-holders tend to react less sensitive to changes in DB's price than customers without a BahnCard – and the same effect can be found for owners of Thalys The Card with respect to Thalys (crosscheck 2). However, these effects are not always statistically significant. Moreover, we do not find price-sensitivity to decrease with the size of the discount granted (crosscheck 3). This indicates that some BahnCard-holders either do not consider the possession of this discount card in their answers or they are unsure whether they would buy a BahnCard in case the examined hypothetical situations became true. A discount card of NS (on the route Cologne-Amsterdam) has no effect on customers' price-sensitivity as it has no connection to any of the presented operators in the scenarios. The own-price effect that is associated with internet-users is either negative or statistically insignificant. This provides some faint evidence that internet-users are more price-sensitive than travellers who buy their tickets offline (hypothesis 7). These results prove fairly robust to changes in the specification of the regression and are not indicative of difficulties like a bias caused by the endogeneity of regressors. In this context, one might argue that certain types of travellers (for example, frequent travellers) are more likely to buy a loyalty card than others. This would cause some of the dependent variables included in our regressions to be in an endogenous causal relationship. We examine this hypothesis and find only weak evidence which would support this objection (see the Appendix). Therefore, we also examine whether this endogeneity impacts the

[7] See $\beta_{2,19}$ in the regression for the route Cologne-Brussels with the data of Thalys-customers, which is statistically insignificant.

coefficients estimated in our multinomial Logit-regressions. We do not find any evidence that the estimated coefficients are subject to a bias caused by the endogeneity of regressors.

Summary of our findings: Our analysis supports hypotheses 1, 2, 3, 4, 6 and 7. We cannot substantiate hypothesis 5. The results of the crosschecks imply that the regression-results are economically plausible.

5 Deriving Choice Probabilities

In this section, we use the above regression to forecast and interpret the choice probabilities of the cooperative and competitive offers for several scenarios. In subsection 5.1, we provide analyses of interviewees' preferences for cooperation versus competition. In subsection 5.2, we analyse interviewees' preferences for the competitors. Subsection 5.3 is concerned with findings regarding travellers' costs of switching among the offers.

5.1 Analysing Interviewees' Choice Probability of Cooperation

The above interpretation of coefficients' signs and significance-levels does not allow for comparisons across the three regressions. This is because the estimated coefficients constitute ratios of the underlying coefficients and the standard deviation of the error term, which differs across the regressions. Therefore, in this section we calculate and compare the probabilities $P_{n,j}$ (see equation (8) in the Appendix) that some homogeneous group of customers n chooses the alternative *rival* (j=1), *Deutsche Bahn* (j=2), or *cooperation* (j=3).

We obtain the aggregate probabilities from these group-specific probabilities by means of aggregation using the classification-method (Ben-Akiva and Lerman 1985: 138). Hence, we approximate the aggregate probability by the sum of group-specific probabilities weighted by the probability ϕ_n that group n is observed among all customers.

$$P_j = \sum_n \phi_n \cdot P_{n,j} \tag{3}$$

In case of the above regressions we consider 512 customer groups, as every of the eight variables $D_{class=1}$, D_{business}, D_{25}, D_{6+}, D_{DB}, $D_{\text{discount\_r}}$, D_{foreign}, and D_{internet} can take a value of 1 or 0. Moreover, BahnCard-customers (D_{DB}=1) may possess one of three types of BahnCard.[8] The probabilities ϕ_n are calculated on basis of the occurrence-probabilities for each of these features in the sample. In case of first-class and business travellers we do not use the sample probabilities but

[8] There are 2^7 combinations of the 7 dummy variables excluding D_{DB}. A traveler may either possess no or one of three types of BahnCard. Therefore, the total number of combinations is $4{\cdot}2^7$=512. In the Thalys-sample we do not observe travelers possessing a BahnCard100. In this sample, the total number of combinations is $3{\cdot}2^7$=384.

probabilities provided by Deutsche Bahn and Thalys which were obtained from surveys larger than our sample. The latter choice affects the calculated values only to a small extent and does not change our qualitative conclusions at all.

Travellers' probability of choosing the alternative *cooperation* is shown on the ordinate in the below graphs for our three samples. On the abscissa we map the price of the entrant relative to the cooperative price (p_e/p_c). We show two scenarios which are possible based on the stated preference data. First, both firms lower their prices by the same amount (see Figure 2). Second, the entrant varies its price while the incumbent leaves its price at the cooperative price (see Figure 3). The alternative *cooperation* is found to be rather meaningless when the two firms set individual prices at a level below 60% of the cooperative price. Therefore, we only show the curves for relative prices above that level.

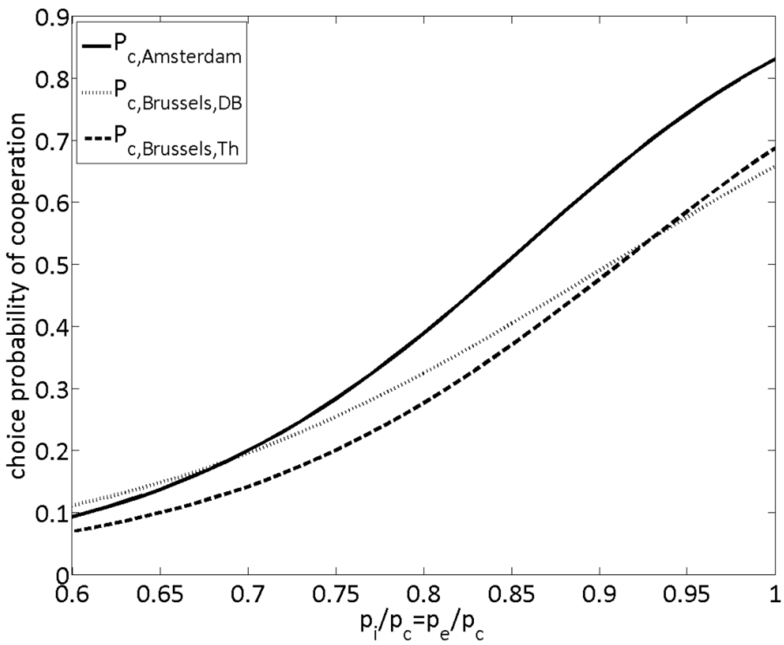


Figure 2: Probability that cooperation is the first choice if entrant and incumbent both lower their price

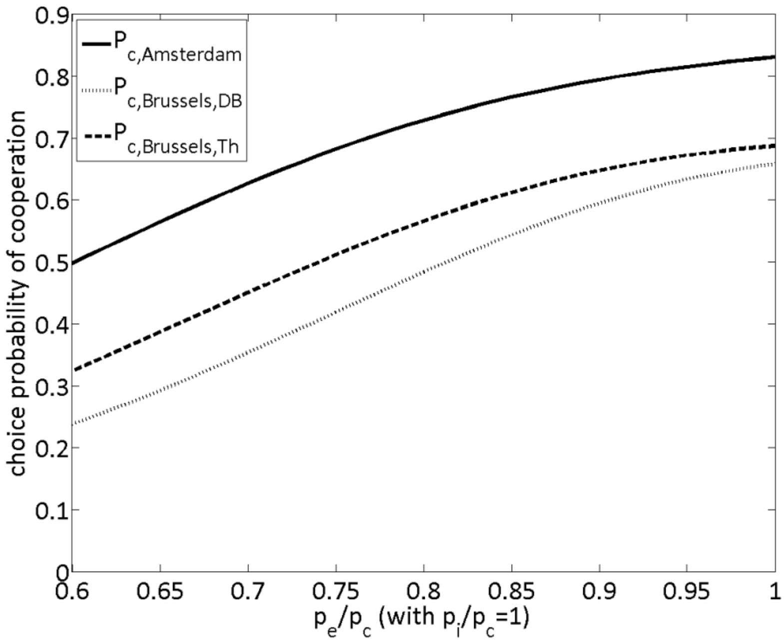


Figure 3: Probability that cooperation is the first choice if only entrant lowers its price

For the situation where the tickets of the incumbent, the entrant and the cooperation among the two are sold at the same price some customers prefer the separate offers to the cooperative one. This feature is present in the raw data and translates into our econometric analysis. In contrast to this observation, we would have expected 100% of respondents to prefer cooperative services if prices are equal, as the service alternative was presented with the highest frequency and with integrated marketing services, i.e. the alternative cooperation should Pareto-dominate the other alternatives. We examined this effect but do not find any evidence that this answering-behaviour is specific to any particular group of customers. Therefore, we provide some possible explanations for this behaviour and suggest further research on this issue:

(a) Some respondents might have a generally positive attitude towards competition and other expected benefits which were not presented in the scenario. In this context, Cantillo et al. (2007: 197) point out that "some individuals may be endowed with a high disposition to change (negative inertia)".

(b) Another explanation might be the prevalence of a status quo bias, i.e. a strong tendency to choose the known operator.

(c) Some groups of customers might not see a value added in the service offered by the cooperation. For example, a private traveller, who is little sensitive to departure times, might already be satisfied with the six departures offered by the incumbent. In combination with the above status quo bias, a customer of this type would choose the incumbent rather than the cooperation at equal prices.

Analysing the curves yields the following conclusions.

1. Most interviewees see benefits in cooperative offers. When the competitors charge identical prices at the level of the cooperative price 66% (Brussels, DB), 69% (Brussels, Thalys), or 83% (Amsterdam) of the interviewed travellers prefer cooperation to one of the competitive offers. When the competitors charge identical prices at a level of 80% of the cooperative price the share of customers that would choose the cooperative offer reduces to 28% (Brussels, Thalys), 32% (Brussels, DB), or 39% (Amsterdam).
2. On the route to Amsterdam we generally find a stronger preference for *cooperation* than on the route to Brussels. First, this finding may indicate that travellers on this route are unfamiliar with competition which is a new, unknown market structure for them. Second, this finding may also reflect that the entrant (RailX) is a hypothetical company so that interviewees' uncertainty about this unknown brand is high. This also results in low choice probabilities for the entrant on the Amsterdam-route.
3. On the route to Brussels we find a relatively higher preference for one of the competitive offers. This might indicate that the travellers on this route are already used to competition and have experienced that the service offered by competitors needs not imply a lower utility than that of an integrated cooperation.

We conclude that the customers have a preference for the market structure and the firms that they know and whose service-quality they can appraise. We thus identify a status quo bias in the choice probabilities.

A majority of the interviewees prefers the cooperative offer to competitive services unless these grant a price discount that is sufficiently high. Our scenario setting implies drawbacks from competition for them. In a situation with infrastructure bottlenecks as assumed in the study, customers for example might have costs from a split of frequencies and from a disintegration of marketing and sales services. They might also have costs from using or learning to use new competitive services. The above choice probabilities with regard to the entrant's price decreases show that a majority of customers expects negative effects from the competitive scenarios presented in the stated preference experiment and requires a price discount to change their preference for cooperation. We can identify the price decreases required to change choice probabilities (see also chapter 5.3) as an indication for the perceived drawbacks form the introduction of competition as presented in our stated preference experiment. It remains to be analysed if benefits from service quality improvements or potential increases in frequency which were not included in our study outweigh these costs of competition.

5.2 Analysing Interviewees' Choice Probabilities for the Entrant

We turn to an analysis of the interviewed passengers' responses when being faced with the decision for one of the two competitors only. Below, we provide the choice probabilities for the entrant on the two routes for different combinations of relative prices. These choice probabilities can be derived from our regressions in the following way. In equations (7) and (8) we define the probability $P_{n,j}$ that some homogeneous group of customers *n* chooses alternative *j*. Therefore, the choice probability of some firm *j* ($\rho_{n,j}$) given that the cooperative offer is not chosen, which occurs with probability 1-$P_{n,c}$, is defined as follows.

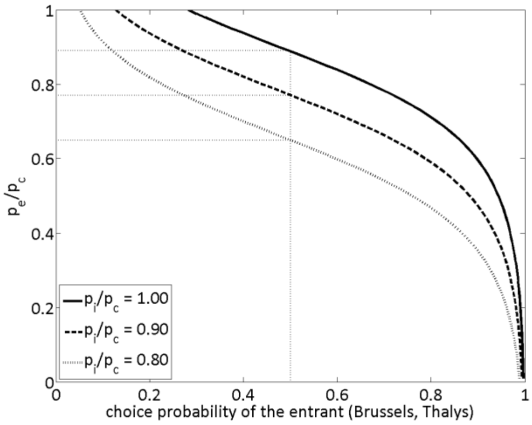
$$\rho_{n,j} = P_{n,j} \big/ \left(1 - P_{n,c}\right) \tag{4}$$

We obtain the aggregate choice probabilities (ρ_j) from the group-specific values by means of aggregation using the classification-method (Ben-Akiva and Lerman 1985: 138) as described above.

In the below graphs, we show the choice probabilities ρ_e of the entrant (i.e. RailX on the Amsterdam-route and DB on the Brussels-route) for varying prices of the entrant. By mapping choice probabilities on the abscissa and prices on the ordinate, these graphs are designed to resemble (residual) demand schedules and, thus, show the (own-)price effects of varying the entrant's price.[9] To get an impression of the effects of changes in the incumbent's price we show the choice probability curves for p_i/p_c equalling 1.0, 0.9, or 0.8. Please note that the curves show the complete pricing range, although very low price ratios and prices of the incumbent below the entrant's prices have not been covered by the scenarios.

The horizontal and vertical dotted lines mark the price that the entrant must charge for an average consumer to be indifferent between choosing the entrant or the incumbent, i.e. both companies are chosen with a probability of 50%. These prices are also shown in the lower part of Table 2 As an example, consider a situation where the incumbent on the Amsterdam route sets its price at the current level (p_i/p_c=1.0). In order to win half of the market the entrant must lower its price to a level of 80% of the current price. Hence, for p_i/p_c=1.0 a traveller's average costs of switching from the incumbent to the entrant amount to 20% of the current price.

[9] The choice probability of the incumbent ρ_i=1-ρ_e can be inferred from these graphs as the horizontal distance between the curves and the vertical axis at the right side.



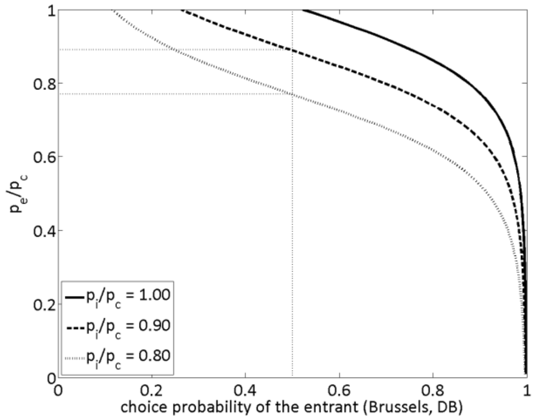


Figure 4: Choice probabilities of the entrant on the route Cologne-Brussels

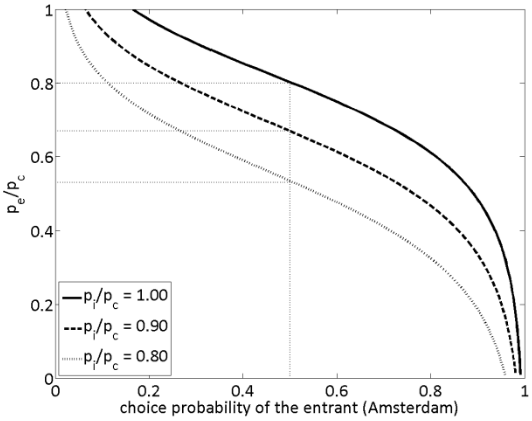


Figure 5: Choice probabilities of the entrant on the route Cologne-Amsterdam

	A'dam	Brussels	
		Thalys	DB
	p_e/p_c	p_e/p_c	p_e/p_c
$p_i/p_c = 1.0$	0.80	0.88	≈ 1.0
$p_i/p_c = 0.9$	0.67	0.77	0.88
$p_i/p_c = 0.8$	0.53	0.65	0.76

Table 2: Price decrease required for 50% choice probability

Analysing the curves yields the following conclusions.

1. The interviewees show a preference for the company on whose trains they were interviewed. This can be seen particularly well for the two Brussels-samples. The travellers who were interviewed on the DB-trains are relatively indifferent between DB and Thalys as long as both companies set about the same prices. The travellers who were interviewed on Thalys-trains are less inclined to choose the entrant DB. In order to make travellers

indifferent between Thalys and DB, the entrant DB must choose lower prices than the incumbent Thalys.

2. Interviewees on the Amsterdam route are reluctant to choose the hypothetical entrant RailX. To make travellers indifferent between choosing the entrant or the incumbent, the entrant must grant a more pronounced price-advantage as in the Brussels case. When the incumbent sets a price at the level of 100% (90%, 80%) of the current price the entrant would have to choose a price at the level of 80% (67%, 53%) of the current price.

The above choice probabilities are aggregated over different groups of customers. This aggregation hides the effect of possessing a discount-card on a firm's market share. Therefore, we calculate the change in the probability for choosing the one or the other firm when a traveller owns a discount or loyalty card. This change is computed relative to a reference group when both rail-companies charge the high cooperative price that is the highest price presented in the scenarios. The reference group consists of second-class leisure travellers, who (i) travelled on this route only once over the last twelve month, (ii) do not possess a discount card, and (iii) did not buy their current ticket on the internet. Regarding the discount cards of Deutsche Bahn and Thalys, we find the following pattern. As stated above, holding a NS-discount card does not have a significant effect.

1. On the Brussels-route, holding a Thalys-card raises the probability of choosing the incumbent Thalys by 25 percentage points (ppt.) in the Thalys-sample and by 28 ppt. in the DB-sample. This means the probability of buying a DB-ticket is reduced by 25 ppt. respectively 28 ppt.
2. On the Brussels-route, buying a BahnCard raises the probability of choosing the entrant DB by 22 ppt. in the Thalys-sample and by 27 ppt. in the DB-sample.
3. On the Amsterdam-route, buying a BahnCard raises the probability of choosing the incumbent DB by 26 ppt.

Our results for Cologne-Amsterdam mirror the risk of a hypothetical bias and its implications for the interpretation of results. The choice probabilities could be used to calculate the market share of an entrant (with additional information on commuters and an extrapolation to represent all customers on the route). Our results imply that an unknown entrant – as on the route Cologne-Amsterdam – should only expect a low market share. Our results for the route Cologne-Brussels indicate that this might only be a temporary phenomenon as customers would get familiar with the new market player resulting in an increase in market share. Therefore, the results for the route Cologne-Amsterdam should not be used to calculate long-term market shares of the entrant. A study for the high speed rail connection Madrid-Barcelona (Román et al. 2007) shows how difficult forecasts on the basis of stated and revealed preference data are. The authors had suggested that even under the least favourable conditions for airlines the rail market share would not exceed 35%

and had questioned the necessity of high speed rail investments between Madrid and Barcelona. However, after two years of operation high-speed rail in 2010 accounted for more than 50% of market share[10] (rail vs. air).

5.3 Findings With Regard to Switching Costs

The regressions confirm two specific cost effects associated with switching rail service providers. Customer card holders show a significant lock-in effect to the issuing company and also reveal a higher preference for cooperation. Internet users appear a bit more price sensitive (direct price effect) and hence have a higher preference for competitive offers than non-internet users.

The main results with regard to costs of changing the service provider are implied in the choice probabilities for cooperation and entrant. They show the overall valuation of the three service alternatives that were presented to the interviewees. For example, we can identify by how much the price of the competitive offers must be lowered in order to gain a certain share from the cooperative service offer. The price discount required to make customers switch then equals the monetary equivalent of their valuation of the disadvantages associated with switching. Due to differences in service frequencies these costs are not fully identical to Klemperer's (1995) definition of "switching cost" which would require functionally equal products. Also these costs would mirror intensifying influences from a status quo bias that might not be related to the switching costs categories. In the following we therefore speak of "changing costs" which include the valuation of all drawbacks perceived by the interviewees in the comparison of the scenario alternatives.

About 17 to 34% of the rail customers see a higher value in the competitive offers if prices of all three alternatives are the same at the level of the current price.[11] All other customers associate a higher value with the cooperative offer and choosing a competitive service is involved with costs for them. For example the below table shows the unilateral price discount that a firm must at least grant before the probability of choosing the cooperative offer drops below 50%. We find that the incumbent must generally lower its price less strongly (i.e. by 10%; Brussels, Thalys) than the entrant (i.e. by 26%; Brussels, Thalys) to attract additional customers from the cooperation. As discussed above, the price discount that any of the two firms on the Amsterdam-route must grant is markedly higher than those required on the Brussels-route.

[10] „High-Speed Rail Gains Traction in Spain." New York Times, March 15 2010. http://www.nytimes.com/2010/03/16/science/earth/16train.html

[11] See Figure 2 and Figure 3.

	A'dam	Brussels	
		Thalys	DB
p_i/p_c	0.17	0.10	0.10
p_e/p_c	0.40	0.26	0.19

Table 3: Minimum unilateral price discount such that the share of cooperation drops below 50%

The different costs of changing to the entrant or to the incumbent indicate the importance of frequencies. It is less costly for rail customers to change from the cooperation with 9 daily train pairs to the incumbent's offer with 6 daily train pairs than to the entrant's offer with only 3 train pairs per day.

We also derive the costs involved with switching between entrant and incumbent from the analysis of preferences for competitors. Without price discount, about 28% of customers (example for Brussels, Thalys[12]) see a higher value in the entrant's service offers. They might have no or negative changing costs. To attract the other 72% of customers the entrant would have to lower its price as they associate a higher value with the incumbent's services. As explained above, we define customers' changing costs as the price discount required to make customers switch. For example, these include the valuation of frequencies and other costs involved with switching as identified by Klemperer (1995). However, we cannot separate the frequency effects from other impacts of service disintegration in marketing (e.g. sales system, information system) which were part of the study. Further research would be required in this field.

The comparison between the three regressions leads us to further indications with regard to switching costs. We identify Klemperer's category "quality of untested brands" in case of RailX on the Amsterdam route: customers are less inclined to choose RailX than to select the entrant on the Brussels route. We have also seen above that rail passengers show a strong tendency to choose the company they were using at the time of the interview. Such a status quo bias is often identified in stated choice experiments. Both rational (switching costs) and psychological persistence effects lead to changing costs.

[12] SeeFigure 5

6 Conclusion

In this paper we present a newly collected dataset of travellers' preferences for long distance passenger rail transport. This dataset was collected by performing almost 700 interviews on-board trains on the two cross-border routes Cologne-Brussels and Cologne-Amsterdam. It contains socio-demographic data of the interviewed travellers, information regarding their trip at the time of the interview and stated-preference data for alternative service scenarios on these routes.

We use this dataset to explore the preferences of the interviewees for different long distance transport services and specifically look for switching costs. This analysis is performed by estimating a multinomial Logit, discrete choice model of demand. As general effects we find that customers with a supposed preference for frequent connections/day (i.e. business- and first class-travellers) have a preference for cooperation among the two competitors, as this allows the travellers to easily use the trains of both firms. Additionally, the regressions support the hypotheses that business travellers react less sensitive to changes in ticket-prices. Travellers, who buy their ticket online react more sensitive to ticket-prices.

We identify considerable costs for switching from cooperation to competitive offers. From a regulatory view our results imply that customers do not automatically benefit from the introduction of competition as they could have costs due to the disintegration of services and a split of frequency. Moreover on the Amsterdam-route, we find that competition after the entry of a new company is a hypothetical situation that customers find hard to imagine. With an unknown entrant they have a high uncertainty about the quality of the entrant. The high uncertainty leads to a high preference for cooperation, while on the Brussels-route we identify a higher preference for the well-known competitive situation. On both routes we find benefits of higher service frequency and of loyalty cards working in favour of cooperation.

With regard to the choice between incumbent and entrant customers tend to prefer the competitor they are using during their current trip which shows a status quo bias based for example on switching costs or habitual travel behaviour. Customers who were interviewed on DB-trains have an especially strong preference for Deutsche Bahn, and customers who were interviewed on Thalys-trains have a relatively stronger preference for Thalys. The unknown entrant on the Amsterdam route faces a high percentage of reluctant customers that rather prefer the known incumbent operator. However, it should be a question of time for customers to get familiar with an entrant's services. We also identify specific switching costs. Customers who possess a discount or loyalty card of one of the service providers have a preference for this firm as compared to its rival and show a lower price sensitivity. So we identify several effects which would work in favour of an incumbent in intramodal choice decisions. In general entrants could be most successful and gain most passengers on routes with a high share of second class leisure travellers that do not possess a customer loyalty card and use the internet for purchasing tickets.

With our stated preference study we analyse demand effects that affect intramodal competition. We look at customer choice behaviour between two rail companies in a simple setting and specifically identify switching costs and psychological persistence behaviour. Our results complement the findings of previous research on competition in passenger rail markets. In particular, taking the status quo bias in intramodal competition into account, entry into commercial passenger rail is more challenging than simulation models so far predicted.

7 References

Banjeree, A. and Summers, L. (1987). “On frequent flyer programs and other loyalty-inducing economic arrangements.” Discussion Paper Number 1337, Harvard Institute of Economic Research, Cambridge

Bates, J. (1988). “Econometric Issues in Stated Preference Analysis.” Journal of Transport Economics and Policy. Vol. 22 No. 1, pp. 59-69

Bates, J. (1998). “Reflections on Stated Preference: Theory and Practice.” In Ortúzar, J. de D., Jara-Diaz, S. and Hensher, D. “Travel Behaviour Research: Updating the State of Play.” Pergamon Press, OxfordBen-Akiva, M. And Lerman, S.R. (1985). “Discrete Choice Analysis: Theory and Application to Travel Demand.” The MIT Press: Cambridge

Ben-Akiva, M. and Lerman, S.R. (1985). “Discrete Choice Analysis: Theory and Application to Travel Demand.” The MIT Press: Cambridge, MA

Bolton, R.N. (1998). “A Dynamic Model of the Duration of the Customer’s Relationship with a Continuous Service Provider: The Role of Satisfaction.” Marketing Science, Vol. 17, pp. 45-65

Bolton, R.N., Kannan, P.K. and Bramlett, M.D. (2000). „Implications of Loyalty Program Membership and Service Experiences for Customer Retention and Value.“ Journal of the Academy of Marketing Science. Vol. 28, No. 1, pp. 95-108

Button, K. (2010). “Transport Economics.” 3rd Edition, Edward Elgar Publishing Limited, Cheltenham

Cameron, A.C. and Trivedi, P.K. (2005). “Microeconometrics Methods and Applications.” Cambridge University Press: Cambridge

Cantillo, V., Ortúzar, J. And Williams, C.W.L (2007). “Modeling Discrete Choices in the Presence of Inertia and Serial Correlation.” Transportation Science, Vol 41, No. 2, pp. 195-205

Carlsson, F. and Löfgren, Å. (2004). “Airline choice, switching costs and frequent flyer programs.” Working Papers in Economics no. 123, Gothenburg University

Dowling, G.R. and Uncles, M. (1997). “Do Customer Loyalty Programs Really Work?” Sloan Management Review. Vol. 38 (4), pp. 71-82

Farrell, J. and Klemperer, P. (2007). “Coordination and Lock-In: Competition with Switching Costs and Network Effects.” In: Armstrong, M. And Porter, R. (Ed.). “Handbook of Industrial Organisation, Volume 3.” Elsevier, pp. 1967-2072

Friederiszick, H., Gantamur, T., Jayaraman, R., Röller, L.-H. and Weinmann,(2009). „Railway Alliances in EC Commercial Passenger Transport: A Competitive Assessment Post-Liberalization 2010." ESMT White Paper No. WP-109-01, ESMT European School of Management and Technology

García-Ferrer, A., Bujosa, M., Juan, A. and Poncela, P. (2006). "Demand Forecast and Elasticities Estimation of Public Transport." Journal of Transport Economics and Policy. Vol. 40 No. 1, pp. 45-67

Gärling, T. and Axhausen, K.W. (2003). „Introduction: Habitual travel choice." Transportation. Vol. 30, pp. 1-11

Goodwin, P.B. (1992). "A Review of New Demand Elasticities with Special Reference to Short and Long Run Effects of Price Changes. " Journal of Transport Economics and Policy. Vol. 26 No. 2, pp. 155-169

Goodwin, P.B. (1977). "Habit and Hysteresis in Mode Choice." Urban Studies, Vol 14, pp. 95-98

Hensher, D.A. and Brewer, A.M. (2001). "Transport – an economics and management perspective." Oxford University Press: Oxford

Hensher, D.A. (2009). "Hypothetical bias, choice experiments and willingness to pay." Working Paper ITLS-WP-09-01, Sydney.

Ivaldi, M., and Vibes, C., (2008). "Price Competition in the Intercity Passenger Transport Market: A Simulation Model." Journal of Transport Economics and Policy, Volume 42, Number 2, pp. 225-254.

Klemperer, P. (1995). "Competition when Consumers have Switching Costs." The Review of Economic Studies, Vol. 62, No. 4, pp. 515-539

Kroes, E.P. and Sheldon, R.J. (1988). "Stated Preference Methods – An Introduction." Journal of Transport Economics and Policy. Vol. 22 No. 1, pp. 11-25

Louviere, J.J. and Hensher, D.A. (2001). "Combining Sources of Preference Data." In Morris, A.E. and Hensher, D.A. "Travel Behaviour Research: The Leading Edge." Pergamon Press, Oxford

Louviere, J.J., Hensher, D.A. and Swait, J.D. (2000). "Stated Choice Methods – Analysis and Application." Cambridge University Press, New York

Oum, T.H., Waters, W.G. and Jong, J.S. (1990). “A Survey of Recent Estimates of Price Elasticities of Demand for Transport.” Working Paper WPS 359, World Bank

Oum, T.H. and Waters, W.G. and Yong J.S. (1992). “Concepts of Price Elasticities of Transport Demand and Recent Empirical Estimates – An Interpretative Survey.” Journal of Transport Economics and Policy. Vol. 26 No. 2, pp. 139-154

Oum, T.H., Gillen, D.W. and Noble, S.E. (1986). “Demand for Fare Class and Pricing in Airline Markets.” Logistics and Transportation Review, 22, pp. 195-222

Preston, J., Whelan, G. and Wardman, M. (1999). “An Analysis of the Potential for On-track Competition in the British Passenger Rail Industry.” Journal of Transport Economics and Policy, 33, Part 1, pp. 77-94

Preston, J. (2009). “Competition for Long Distance Passenger Rail Services: The Emerging Evidence.” OECD/ITF, Discussion Paper No. 2009-23

Román, C., Espino, R. and Martín, J.C. (2007). “Competition of high-speed train with air transport: The case of Madrid-Barcelona.” Journal of Air Transport Management, 13, pp. 277-284

Samuelson, W. and Zeckhauser, R. (1988). “Status Quo Bias in Decision Making.” Journal of Risk and Uncertainty, Vol. 1, pp. 7-59

Seabright, P., et al. (2003). “The Economics of Passenger Rail Transport: A Survey.” IDEI Report No. 1 on Passenger Rail Transport

Steer Davis Gleave (2004) “EU Passenger Rail Liberalisation: Extended Impact Assessment.” Study for European Commission, DG-TREN

Train, K.E. (2009). “Discrete Choice Methods with Simulation.” 2nd edition. Cambridge University Press: Cambridge

Verhoef, P.C. (2003). “Understanding the Effect of Customer Relationship Management Efforts on Customer Retention and Customer Share Development.” Journal of Marketing, Vol. 67 (October 2003), pp. 30-45

Whitaker, B., Terzis, G., Soong, E. and Yeh, W. (2005) “Stated Preference as a Tool to Evaluate Airline Passenger Preferences and Priorities.” Transportation Research Record: Journal of the Transportation Research Board, No. 1915, Transportation Research Board of the National Academies, Wahington D.C., pp. 55-61

Widlert, S. (1998). “Stated Preference Studies: The Design Affects the Results.” In Ortúzar, J. de D., Jara-Diaz, S. and Hensher, D. “Travel Behaviour Research: Updating the State of Play.” Pergamon Press, Oxford

8 Appendix

In this appendix, we present a brief review of discrete choice modelling and show the results of our regressions. Moreover, we present results of some tests which indicate that the coefficients calculated by means of these regressions are unlikely to suffer from an endogeneity bias. Furthermore, we do not find evidence that a possible correlation of an interviewee's answers across the 21 scenarios biases our results.

As presented in section 4, we assume that some decision maker *n* obtains utility U_{nj} from alternative *j* with *j*=1 denoting the *rival* of Deutsche Bahn, *j*=2 denoting *Deutsche Bahn*, and *j*=3 applying for the *cooperation* offer. This utility is unobservable. However, in our dataset we observe the ranking order of the three alternatives as chosen by the interviewee. A rational decision maker prefers alternative *i* to alternative *j* if and only if $U_{ni}>U_{nj}\ \forall\ j\neq i$. The basic idea of discrete choice modelling is to infer information about the interviewees' unobservable decision making process by analysing their observable choices.

We do this by econometrically relating interviewees' choices to observable attributes of (i) the three above alternatives *j* and of (ii) the decision maker. These attributes, labelled $x_{nj}\ \forall\ j$, are described in section 3 and summarised in Table 4. We specify a representative utility function that relates these observed factors to the decision maker's observed utility

$$V_{nj} = \boldsymbol{\beta}_j \,' \boldsymbol{x}_{nj} \forall j \,, \tag{4}$$

where $\boldsymbol{x}_{nj}$ is a vector of variables that relate to alternative *j* as faced by decision maker *n*. $\boldsymbol{\beta}_j$ denotes the coefficients of these variables. The specific functional form for the observed part of utility of this problem is described by equation (2). Since there are aspects of utility that the researcher does not or cannot observe, the observed part of utility V_{nj} does not perfectly equal the true utility U_{nj}. Utility is decomposed as

$$U_{nj} = V_{nj} + \varepsilon_{nj} \,, \tag{5}$$

where ε_{nj} captures the factors that affect utility but are not included in V_{nj}. The researcher does not know $\varepsilon_{nj}\ \forall\ j$ and therefore treats these terms as random.

	0	**1**
$\boldsymbol{D}_{\text{class=1}}$	2nd class	1st class
$\boldsymbol{D}_{\text{business}}$	private	business travel
$\boldsymbol{D}_{25}$	otherwise	2-5 rail travels on this route/year
$\boldsymbol{D}_{6+}$	otherwise	6 or more rail travels on this route/year
$\boldsymbol{D}_{\text{DB}}$	otherwise	possession of a BahnCard
$\boldsymbol{D}_{\text{DB25}}$	otherwise	possession of a BahnCard 25
$\boldsymbol{D}_{\text{DB50}}$	otherwise	possession of a BahnCard 50
$\boldsymbol{D}_{\text{DB100}}$	otherwise	possession of a BahnCard 100
$\boldsymbol{D}_{\text{discount\_r}}$	otherwise	possession of a RailPlus Card (NS)
		or: possession of Thalys The Card
$\boldsymbol{D}_{\text{internet}}$	otherwise	ticket was bought online
$\boldsymbol{D}_{\text{direct}}$	otherwise	start/end of current trip is no further than
		Cologne/Amsterdam
		or: Cologne/Brussels

Table 4: Variable definitions

The joint density of the random vector $\varepsilon'_n = \langle \varepsilon_{n1}, \ldots, \varepsilon_{nJ} \rangle$ is denoted $f(\varepsilon_n)$. With this density, the researcher can make probabilistic statements about the decision maker's choice. The probability that decision maker *n* chooses alternative *i* is

$$\begin{aligned} P_{ni} &= \text{Prob}(U_{ni} > U_{nj} \forall j \neq i) \\ &= \text{Prob}(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \forall j \neq i) \\ &= \int_{\varepsilon} I(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \forall j \neq i) f(\varepsilon_n) d\varepsilon_n \end{aligned} \tag{6}$$

where $I(\cdot)$ is the indicator function, equalling 1 when the expression in parentheses is true and 0 otherwise (Train 2009: 15). The density $f(\varepsilon_n)$ is the distribution of the unobserved portion of utility within the population of people who face the same observed portion of utility (see our analysis in section 5). Given the assumption that each ε_{nj} is independently, identically distributed extreme value we can estimate V_{nj} by means of a multinomial logit model and compute the choice probability P_{ni} of alternative *i* by decision maker *n* as

$$P_{ni} = \frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}}. \tag{7}$$

Based on this theoretical groundwork of our empirical analysis we specify the interviewees' utility function as described in section 4.1. Concerning the interpretation of the below regression tables note that the same variables enter the utility-function of all alternatives. Therefore, any model with the same difference in corresponding coefficients is equivalent (Train 2009: 20) so that the utility and the coefficients of the alternative *cooperation* are normalized to zero. As a consequence the estimated coefficients for the alternatives *rival* and *Deutsche Bahn* must be interpreted relative to the alternative *cooperation*.

In the following, we provide the regressions underlying the evaluations presented in the main part of this paper. The column E(sign) shows the expected sign of the estimated coefficients according to the hypotheses provided in section 4.1 For example, (-) / (+) means that the coefficient associated with the *rival* is expected to be negative, while that associated with *Deutsche Bahn* is expected to be positive. (-) means that both coefficients are expected to be negative. (±) means that we do not have particular expectations regarding the sign of the respective coefficient. *, **, or *** denote the 10%, 5%, or 1% level of significance. The variables used must be interpreted as shown in Table 4. Prices of Deutsche Bahn (p_{DB}) and its rival (p_r) are defined as percentages of the cooperative price (p_c).

			Cologne to Amsterdam					
			observed utility (rival, j=1)			observed utility (DB, j=2)		
		E(sign)	coefficient	std. error		coefficient	std. error	
$\beta_{j,1}$	constant	(±)	8.8428	0.6615	***	11.3020	0.6227	***
$\beta_{j,2}$	$D_{\text{class=1}}$	(-)	-1.5973	0.2142	***	-1.2394	0.1678	***
$\beta_{j,3}$	D_{business}	(-)	-6.6087	0.8142	***	-4.4295	0.6418	***
$\beta_{j,4}$	D_{25}	(±)	-1.5013	0.1425	***	-0.4107	0.1311	***
$\beta_{j,5}$	D_{6+}	(±)	-0.8763	0.2241	***	-0.2205	0.1721	
$\beta_{j,6}$	D_{DB}	(-)	-2.3220	0.7484	***	-1.4570	0.5719	**
$\beta_{j,7}$	$D_{\text{discount\_r}}$	(-)	-0.3287	0.7499		0.6876	0.6981	
$\beta_{j,9}$	p_r/p_c	(-) | (±)	-9.0442	0.7223	***	-1.5133	0.6677	**
$\beta_{j,10}$	$(p_r/p_c)\cdot D_{\text{discount\_r}}$	(+) | (±)	0.7858	1.1917		-0.4154	1.0339	
$\beta_{j,12}$	$(p_r/p_c)\cdot D_{\text{business}}$	(+) | (±)	3.4270	0.8851	***	1.1820	0.6952	*
$\beta_{j,13}$	$(p_r/p_c)\cdot D_{\text{internet}}$	(-) | (±)	-1.3391	0.8091	*	0.8478	0.6977	
$\beta_{j,14}$	p_{DB}/p_c	(±) | (-)	-1.7327	0.7568	**	-11.1562	0.7875	***
$\beta_{j,16}$	$(p_{DB}/p_c)\cdot D_{\text{DB25}}$	(±) | (+)	1.3230	0.8935		1.8962	0.7161	***
$\beta_{j,17}$	$(p_{DB}/p_c)\cdot D_{\text{DB50}}$	(±) | (+)	0.8445	0.8894		1.6879	0.7153	**
$\beta_{j,18}$	$(p_{DB}/p_c)\cdot D_{\text{DB100}}$	(±) | (+)	0.2290	0.9392		0.7006	0.7612	
$\beta_{j,19}$	$(p_{DB}/p_c)\cdot D_{\text{business}}$	(±) | (+)	2.5872	0.9359	***	3.0619	0.8540	***
$\beta_{j,20}$	$(p_{DB}/p_c)\cdot D_{\text{internet}}$	(±) | (-)	1.4461	0.5652	**	-0.5377	0.5587	
$\beta_{j,21}$	D_{direct}	(+)	0.5470	0.6754		0.5483	0.5784	
$\beta_{j,22}$	$(p_r/p_c)\cdot D_{\text{direct}}$	(-) | (±)	-1.3691	0.7955	*	1.0522	0.6911	
$\beta_{j,23}$	$(p_{DB}/p_c)\cdot D_{\text{direct}}$	(±) | (-)	0.3184	0.8189		-1.8876	0.8262	**
Mc-Fadden R^2		36.54%						
number of observations		3,780						

Table 5: Amsterdam - Choice among all three alternatives

			Cologne to Brussels (DB trains)					
			observed utility (Thalys, *j* =1)			observed utility (DB, *j* =2)		
		E(sign)	coefficient	std. error		coefficient	std. error	
$\beta_{j,1}$	constant	(±)	7.7507	0.4338	***	5.3160	0.4015	***
$\beta_{j,2}$	$D_{\text{class=1}}$	(-)	-0.0659	0.1375		-0.7687	0.1167	***
$\beta_{j,3}$	D_{business}	(-)	-1.9052	0.5157	***	-1.6665	0.4731	***
$\beta_{j,4}$	D_{25}	(±)	0.0984	0.1056		0.2037	0.0880	**
$\beta_{j,5}$	D_{6+}	(±)	-0.3084	0.1441	**	0.6880	0.1072	***
$\beta_{j,6}$	D_{DB}	(-)	-1.6957	0.3241	***	-0.6507	0.2514	***
$\beta_{j,7}$	$D_{\text{discount\_r}}$	(-)	-4.9809	2.1133	**	-5.0551	2.3651	**
$\beta_{j,9}$	p_r/p_c	(-) | (±)	-11.5444	0.7463	***	0.2248	0.5040	
$\beta_{j,10}$	$(p_r/p_c)\cdot D_{\text{discount\_r}}$	(+) | (±)	6.5467	2.6505	**	5.8805	2.8543	**
$\beta_{j,12}$	$(p_r/p_c)\cdot D_{\text{business}}$	(+) | (±)	3.7570	0.8894	***	-0.7248	0.6083	
$\beta_{j,13}$	$(p_r/p_c)\cdot D_{\text{internet}}$	(-) | (±)	-1.3030	0.8346		1.5066	0.4674	***
$\beta_{j,14}$	p_{DB}/p_c	(±) | (-)	2.5280	0.7021	***	-7.2831	0.4856	***
$\beta_{j,16}$	$(p_{DB}/p_c)\cdot D_{\text{DB25}}$	(±) | (+)	1.8392	0.4903	***	1.6117	0.3851	***
$\beta_{j,17}$	$(p_{DB}/p_c)\cdot D_{\text{DB50}}$	(±) | (+)	1.2034	0.4974	**	1.1124	0.3852	***
$\beta_{j,18}$	$(p_{DB}/p_c)\cdot D_{\text{DB100}}$	(±) | (+)	-0.5387	0.9140		2.5271	0.4842	***
$\beta_{j,19}$	$(p_{DB}/p_c)\cdot D_{\text{business}}$	(±) | (+)	-2.4289	0.8236	***	2.0086	0.5479	***
$\beta_{j,20}$	$(p_{DB}/p_c)\cdot D_{\text{internet}}$	(±) | (-)	1.5008	0.9458		-1.6874	0.5927	***
$\beta_{j,21}$	D_{direct}	(+)	1.9485	0.9875	**	2.7787	0.9317	***
$\beta_{j,22}$	$(p_r/p_c)\cdot D_{\text{direct}}$	(-) | (±)	-2.9912	1.6258	*	-1.1561	1.1760	
$\beta_{j,23}$	$(p_{DB}/p_c)\cdot D_{\text{direct}}$	(±) | (-)	1.4276	1.4075		-1.3137	0.9534	
Mc-Fadden R^2		23.92%						
number of observations		5,116						

Table 6: Brussels (DB) - Choice among all three alternatives

			Cologne to Brussels (Thalys trains)					
			observed utility (Thalys, j=1)			observed utility (DB, j=2)		
		E(sign)	coefficient	std. error		coefficient	std. error	
$\beta_{j,1}$	constant	(±)	10.1105	0.5142	***	6.9851	0.5043	***
$\beta_{j,2}$	$D_{class=1}$	(-)	-0.3372	0.1674	**	-0.9523	0.1847	***
$\beta_{j,3}$	$D_{business}$	(-)	-3.2024	0.6610	***	-2.5943	0.7191	***
$\beta_{j,4}$	D_{25}	(±)	-0.4852	0.1159	***	0.1488	0.1095	
$\beta_{j,5}$	D_{6+}	(±)	0.6249	0.1635	***	0.6741	0.1706	***
$\beta_{j,6}$	D_{DB}	(-)	-0.0165	0.3279		0.2753	0.3199	
$\beta_{j,7}$	$D_{discount\_r}$	(-)	-4.3963	1.0012	***	-3.9530	1.5692	**
$\beta_{j,9}$	p_r/p_c	(-) | (±)	-12.8689	0.7455	***	-1.5754	0.6148	***
$\beta_{j,10}$	$(p_r/p_c)\cdot D_{discount\_r}$	(+) | (±)	4.6170	1.2136	***	2.2332	1.8458	
$\beta_{j,12}$	$(p_r/p_c)\cdot D_{business}$	(+) | (±)	5.0164	0.9204	***	0.9093	0.8432	
$\beta_{j,13}$	$(p_r/p_c)\cdot D_{internet}$	(-) | (±)	-1.2294	0.7152	*	-0.5592	0.5636	
$\beta_{j,14}$	p_{DB}/p_c	(±) | (-)	1.7635	0.6218	***	-7.3740	0.5255	***
$\beta_{j,16}$	$(p_{DB}/p_c)\cdot D_{DB25}$	(±) | (+)	-1.4504	0.5604	***	-0.8157	0.5509	
$\beta_{j,17}$	$(p_{DB}/p_c)\cdot D_{DB50}$	(±) | (+)	-0.5215	0.5153		0.1299	0.5072	
$\beta_{j,19}$	$(p_{DB}/p_c)\cdot D_{business}$	(±) | (+)	-1.9451	0.7704	**	0.7768	0.7774	
$\beta_{j,20}$	$(p_{DB}/p_c)\cdot D_{internet}$	(±) | (-)	1.4431	0.8541	*	0.5984	0.7525	
$\beta_{j,21}$	D_{direct}	(+)	-1.2577	0.7101	*	-0.8455	0.7507	
$\beta_{j,22}$	$(p_r/p_c)\cdot D_{direct}$	(-) | (±)	2.0179	1.0090	**	1.3303	0.9196	
$\beta_{j,23}$	$(p_{DB}/p_c)\cdot D_{direct}$	(±) | (-)	-0.5052	0.8610		-0.4187	0.8297	
Mc-Fadden R^2		27.11%						
number of observations		3,675						

Table 7: Brussels (Thalys) - Choice among all three alternatives

The above multinomial Logit-regressions include variables that may be subject to endogeneity issues. In particular, the choice to buy a BahnCard or the loyalty card of another operator might depend on the number of travels on this route during the last year (*f*), the state of being a business traveller ($D_{business}$=1), and being a traveller who buys his/her tickets over the internet ($D_{internet}$=1). These relationships are shown in Table 7 where we present the results of multinomial Logit-regressions for choosing one of the three types of the BahnCard and a binary Logit-regression for choosing the loyalty card of another rail-company. In the dataset collected from Thalys-customers at the Brussels-route, the multinomial Logit-regression does not include the BahnCard100-alternative as customers with such a loyalty card are not observed in this sample.

			Multinomial Logit										**Binary Logit**		
			BahnCard 25 (*k*=1)			**BahnCard 50 (*k*=2)**			**BahnCard 100 (*k*=3)**				**Loyalty Card Rival (*l*=1)**		
			coefficient	std. error		coefficient	std. error		coefficient	std. error			coefficient	std. error	
Amsterdam	frequency (f)	$\gamma_{k,1}$	-0.9778	1.0641		-1.1823	0.9587		-2.7226	1.2683	**	$\delta_{l,1}$	2.3439	1.1981	**
	f^2	$\gamma_{k,2}$	0.2757	0.1926		0.3715	0.1705	**	0.5945	0.2203	***	$\delta_{l,2}$	-0.3033	0.1982	
	$D_{business}$	$\gamma_{k,3}$	1.1068	0.5717	*	0.7014	0.5041		0.5993	0.7448		$\delta_{l,3}$	-1.6722	0.6661	**
	$D_{internet}$	$\gamma_{k,4}$	1.6817	0.5503	***	1.2270	0.4918	**	0.9030	0.7142		$\delta_{l,4}$	-0.9709	0.6742	
	constant	$\gamma_{k,5}$	-2.5763	1.3269	*	-2.0118	1.2158	*	-0.7102	1.4898		$\delta_{l,5}$	-4.9682	1.6468	***
		Mc-Fadden R^2		15.86%								Mc-Fadden R^2		12.31%	
		no. of obs.		180								no. of obs.		180	
Brussels (DB)	frequency (f)	$\gamma_{k,1}$	-0.2509	0.5954		0.6101	0.5718		1.2431	1.3871		$\delta_{l,1}$	-2.917935	2.641759	
	f^2	$\gamma_{k,2}$	0.0952	0.1046		-0.0249	0.0986		-0.1509	0.2343		$\delta_{l,2}$	0.5597532	0.4557545	
	$D_{business}$	$\gamma_{k,3}$	-0.0016	0.3420		0.4064	0.3306		0.3943	0.8080		$\delta_{l,3}$	-0.243706	1.466779	
	$D_{internet}$	$\gamma_{k,4}$	0.4062	0.3573		0.1926	0.3571		-14.2236	756.7928		$\delta_{l,4}$	0.9848289	1.443586	
	constant	$\gamma_{k,5}$	-1.1614	0.6579	*	-2.2529	0.6815	***	-4.5257	1.7126	***	$\delta_{l,5}$	-3.041708	2.621983	
		Mc-Fadden R^2		5.39%								Mc-Fadden R^2		11.25%	
		no. of obs.		248								no. of obs.		248	
Brussels (Thalys)	frequency (f)	$\gamma_{k,1}$	0.3853	0.2138		-0.0843	0.2412		not applicable			$\delta_{l,1}$	0.8305	0.3271	**
	f^2	$\gamma_{k,2}$	-2.1341	1.0656		0.1897	0.5216					$\delta_{l,2}$	0.2118	0.8364	
	$D_{business}$	$\gamma_{k,3}$	-0.8164	0.7804		0.3735	0.5358					$\delta_{l,3}$	0.2208	0.7874	
	$D_{internet}$	$\gamma_{k,4}$	-0.1891	0.7139		-0.4138	0.5941					$\delta_{l,4}$	1.3657	0.7793	*
	constant	$\gamma_{k,5}$	-2.6362	0.5957	***	-2.0153	0.5267	***				$\delta_{l,5}$	-5.9182	1.3393	***
		Mc-Fadden R^2		4.91%								Mc-Fadden R^2		16.72%	
		no. of obs.		183								no. of obs.		183	
Definition	f =	1	= 1												
		2	= 2-3												
		3	= 4-5												
		4	= 6-12												
		5	> 12 journeys on this route during the last 12 months												

Table 8: Determinants of choosing a loyalty card

These regressions show that there are only weak relationships between the regressors included in our main multinomial Logit-regressions. We do not find any indication that these effects cause a bias in the coefficients estimated by these regressions. If the endogeneity between the loyalty card-variables and the other regressors had a biasing impact on the estimated coefficients we would expect to see a perceptible change in the estimated coefficients when excluding either of these two sets of independent variables. Hence, we exclude the loyalty card-variables and re-run the regressions for the choice of either rail-service provider. In all of the three regressions, all the remaining coefficients remain at vastly the same values as in the previous case. Similarly, the values of the Pseudo-R^2 drop only slightly from 27.11% to 26.18% (Brussels, Thalys), 23.92% to 21.92% (Brussels, DB), and 36.54% to 34.42% (Amsterdam). Therefore, we do not find any indication that the endogeneity of regressors would have caused a bias in the estimated coefficients.

Besides the issue of endogeneity, the results of our above regression may also be affected by heterogeneity across the answers of respondents in each group of customers. This may also pose difficulties because in our stated preference data with 21 scenarios each respondent accounts for 21 observations that may potentially be correlated. Therefore, we estimate a mixed logit model with random coefficients for the dummy variables ($\beta_{j,2}$ to $\beta_{j,7}$ and $\beta_{j,21}$) and the own-price effects ($\beta_{1,9}$ and $\beta_{2,14}$). Based on a likelihood ratio test we reject the hypothesis that a model with fixed coefficients as used in the main part of this paper describes the data as well as a mixed logit model with random coefficients. Nonetheless, for the following reasons we decide to use the model with fixed coefficients rather than the one with random coefficients.

First, we do not find evidence that our qualitative results (see section 4.2) differ across these models. Second, we find that the in-sample choice probabilities (see section 5.1) predicted by the two models are the same in expectation. We plot a histogram for each alternative *j* of the difference between the choice probabilities calculated from the two models for each observed individual and scenario. These differences follow bell-shaped distribution functions with mean zero. Third, we consider the results of the above multinomial logit regressions with fixed coefficients more reliable than those obtained with random coefficients. This is because the *mlogit*-algorithm, which we use in Stata to obtain the multinomial logit results, returns very reliably to the presented results when we run the algorithm multiple times and requires a runtime of only few seconds. The *mixlogit*-algorithm, which we use, requires about 1.5 hours on a desktop computer to perform the random coefficient regression at default precision and returns different results when we repeat the calculation several times. Given that our datasets are large and we consider many explanatory variables this problem is only somewhat ameliorated when we raise the precision of the calculation.