The impact of fare complexity on rail demand

Paulo Rui Anciaes
Centre for Transport Studies, University College London

Paul Metcalfe
PJM Economics

Chris Heywood
Rob Sheldon
Accent

Abstract

The current rail fare structure in the UK is widely considered to be complex. It offers flexibility by including many different types of tickets but it can also cause confusion, which may lead some to a negative view towards rail travel and thereby potentially inhibit demand, in comparison to a simpler fare regime. This study used an innovative stated preference survey to quantify the demand effects of fare complexity, focusing on Advance tickets (those that are restricted to a particular train service). The choice experiment was designed to mirror very closely the actual booking experience when buying tickets online, in all its complexity. Participants could choose among up to 531 different ticket type combinations for the outward and return legs of a trip and from up to 25 possible train services for each leg. The key design attribute was complexity, defined as the range of different Advance tickets on offer. The survey was applied to a sample of 1,027 users and 179 non-users of the rail network on the London-Leeds route. The modelling of the choices with a nested mixed logit model suggested that, all else equal, reducing complexity by removing Advance tickets would lead to a substantial reduction of demand (11% to 45%, depending on route segment). Equalizing the price of Advance tickets for all train services was predicted to cause a smaller reduction (3-6%). By contrast, increasing complexity by adding new Flexible Advance tickets (valid on the services immediately before or after the chosen service) would increase demand by 4-15%. These findings run counter to the hypothesis that simplifying the fare structure would lead to increases in demand for rail travel.

1. Introduction

A complex rail fare structure brings potential benefits for customers, insofar as it offers greater opportunity to find the preferred combination of cost, duration, comfort, and flexibility. But complexity also increases the time and effort to evaluate all the alternatives, which may influence perceptions about rail travel and willingness to travel by train. Experiments in the fields of psychology and consumer behaviour suggest that complex choice situations have both positive (Kahn and Lehmann 1991, Reibstein *et al.* 1975) and negative impacts (Swait and Adamowicz 2011, Ariely 2000, Garbarino and Edell 1997, Iyengar and Lepper 2000) on attitudes towards the choice, but there is little quantitative evidence on the impact of fare complexity on the demand for rail travel. In urban areas, simplifying the fares of multi-modal commuting trips usually leads to increased demand. However, this does not necessarily apply to interurban rail transport, which has a more varied demand and supply than urban public transport.

Understanding the impact of fare complexity on demand is particularly important in the UK, where the variety of interurban rail fares has increased dramatically since privatisation in the 1990s. The current fare regime at the national level includes four broad types of tickets: "Anytime" (with no time restrictions), "Off-Peak" (not valid at peak times), "Super Off-Peak" (only valid at the quietest times of the day) and "Advance" (booked before the trip and valid only on a specific service). Anytime, Off-Peak, and Super Off-Peak tickets can be "Single" (valid for one trip leg), "Day Single" (valid on a specific day), "Return" (valid for two legs of a trip), or "Day Return" (valid for two legs of a trip on a specific day). Most tickets are available in "First" or "Standard" class. There are also season tickets, valid for more than one trip. Some of the fares (including all Advance fares) are unregulated, so they differ in price and in the type of restrictions for each of the train-operating companies (TOCs) serving the same route. However, some tickets can be used in the services of more than one company. The system also has a number of anomalies, such as the possibility of making savings by splitting tickets, i.e. using more than one ticket for separate sections of the same single trip.

The complexity created by the variety of ticket types on offer is compounded by problems specific to each retail outlet. Ticket machines do not always show all the available options (Passenger Focus 2010, ESA Retail 2018) and the experiences in ticket offices vary with the knowledge of the staff (Which? 2011a). Online booking systems usually show most available ticket types and information about their restrictions, but by doing so, they also increase the effort that the user must make to find the most suitable ticket (Which? 2011b). The tendency

for an increase in online sales (SDG 2016), together with an increase in the share of individual tickets, rather than season tickets, due to more flexible working hours (KPMG 2018), have increased the need to understand how people perceive the fare system and the different ticket types.

Previous research in the UK has found that people do value the existence of a range of different ticket types (BDRC Continental 2015). However, there is also a general perception that the fare structure is "complicated, confusing, and illogical" (Passenger Focus 2012, p.8). Many people are not aware of the restrictions associated with each ticket type, including the exact hours for Off-Peak and Super Off-Peak tickets (which vary by direction of travel) and the train services that can be used with Advance tickets (including the possibility of using stations, routes, and trains other than those specified in the ticket) (Which? 2011b). As a result, passengers often buy tickets that are not the most advantageous for them (SPAFT 2011, ORR 2012). This problem is also relevant beyond the UK. For example, a European-wide survey of rail stakeholders has found issues about fare complexity in 13 out of 20 countries (SDG 2016, p.78), a problem that may be amplified by the upcoming deregulation of the passenger rail market in the European Union. However, despite this evidence, no studies have investigated whether negative perceptions about fare complexity actually lead to a lower willingness to choose rail travel.

The objective of the present study is to understand the impacts of potential changes to the rail fare structure in the UK on passenger rail demand, which has implications for similar policies in other sectors and countries. Complexity is defined as the range of different ticket types, with different prices and restrictions. The study focuses on the impact of changing Advance tickets, because of their wide variety and because they are the ticket types with more restrictions in terms of departure time and train-operating company, and so, they are the most difficult to understand. We look at the impact of scenarios that simplify the current structure (by equalizing the price of Advance tickets for all train services, or removing Advance tickets altogether) or add extra complexity (by introducing new types of Advance tickets).

The study also makes a methodological contribution to research on complex choices made online. We develop an innovative stated preference survey of users and non-users of the rail network to elicit choices regarding ticket types and train services under different fare regimes, using a survey questionnaire that replicates a real-work online booking engine. The impact of complexity on demand is estimated with a nested choice model that controls for factors other

than complexity that also influence choices, such as fare levels, ticket restrictions, trip duration, deviations from ideal departure or arrival time, and company used.

The paper proceeds as follows. Section 2 is a review of the literature on the impact of complexity of fare structures on rail demand. Section 3 describes the survey design and sample characteristics. Section 4 presents results on general perceptions regarding fare structures and ticket types. Section 5 describes the design of the stated preference exercise and Section 6 presents models of the choices of ticket types and train services. Section 7 estimates the demand impact of different fare complexity scenarios in key route segments. Section 8 synthesizes the results and discusses their implications for policy and research.

2. Literature review

The move towards liberalization of rail markets in many countries has opened the possibility for train-operating companies to differentiate their products in order to compete with each other and to use yield management and price discrimination to maximize revenue and better allocate customers to services. These strategies have been facilitated by the gradual shift to online sales, which reduces the companies' "menu costs" i.e. the costs to change prices. In the UK, these trends have led to a greater variety of ticket types on offer, as companies exploited differentiation in features such as time of purchase (in advance or on the day of travel) and time of travel (peak or off-peak). This type of competition tends to result in substitution effects in the demand for the services provided by different companies (Preston 1999, Preston 2008, Wardman and Toner 2003, Whelan *et al.* 2008), although there is also resistance to changing to a new company (Paha *et al.* 2013). But the impact of competition through ticket types on overall rail demand is still unclear, as past studies have not always separated substitution from scale effects, and when they did, they have not isolated the effect of the range of different ticket types available from the effects of the attributes of each ticket type.

This issue is relevant because the number of alternatives available to customers increases the complexity of the choice, which can be perceived positively or negatively – a question that has been studied for decades in the fields of psychology and consumer behaviour. In general, individuals like to have many options to choose from, as this offers flexibility and a greater chance of finding a product that matches preferences (Kahn and Lehmann 1991). They also attach "option value" to large and varied choice sets and tend to consume more when faced with those choice sets (Reibstein *et al.* 1975). But in many contexts, they dislike complex choices. For example, participants in the study by Iyengar and Lepper (2000)

showed more motivation to make choices and satisfaction with those choices when the set of options was limited. In the case of rail travel, London Economics (2015) found that passengers are less confident about their choices when too much information is provided at once.

There are two main reasons for this negative perception of complexity. The first reason is that choices with many alternatives have higher "search costs" (the time and effort put into finding the best option) (Swait and Adamowicz 2001, Lee and Lee 2004). The second reason is that complex choice sets increase the risk of choosing the wrong option (Ariely 2000), so some people may prefer a situation where they understand all the options, rather than one that potentially provides higher utility but where they do not understand the options (Ellsberg 1961, Garbarino and Edell 1997). In the case of the transport market, the main issues are the uncertainty of the passengers about the correct fare and the concern that they will pay more than they need to. For example, Hess *et al.* (2013) found that people's willingness to pay to have fixed or zonal systems is lower when accounting for the uncertainty about the current fare and the number of zones.

Complexity may therefore limit the demand for public transport. However, the evidence confirming this hypothesis comes mostly from cases where there was simplification, i.e. a reduction of complexity. Around the world, policies to simplify public transport fares (through travel cards, flat fares, zone systems, or integration of services offered by different providers) have usually increased demand and revenue, besides improvements in passenger satisfaction, boarding times, detection of fraud, and transaction and administration costs (Booz & Co 2009). Academic studies have also found a positive link between demand and fare simplification in multimodal urban transport systems (Taylor and Carter 1998, Matas 2004) and specific modes like buses (Sharaby and Shiftan 2012) and taxis (Castillo-Manzano and Sánchez-Braza 2011). Few studies have looked at interurban passenger transport and the ones that did concluded that the demand impact of fare integration is moderate (Abrate *et al.* 2009) or depends on price and income (Dargay and Pekkarinen 1997).

It is also difficult to pinpoint the main factor driving the demand impacts of changes in the fare structure, because those impacts can be due to changes in complexity or to the associated changes in fare levels. For example, increasing complexity by differentiating fares by time of day may lead to a reduction in overall rail travel because the fare at some times of the day becomes higher (Weesie *et al.* 2009). Conversely, simplifying the fare structure by integrating tickets of different companies usually means that passengers do not have to pay extra for

transfer between different services, which reduces the cost of using public transport (Sharaby and Shiftan 2012). Notwithstanding, Taylor and Carter (1998) found that fare simplification had a positive impact on demand, even in a context of a general fare increase.

It is likely that people balance the advantages of having a high number of options with the effort required to understand the differences between those options. However, there is almost no empirical evidence on how this balance is achieved in the case of interurban rail travel. The conclusions of studies that focused on urban public transport systems may not apply to the case of interurban rail transport, whose customers tend to have a different mix of trip purposes, and higher flexibility in departure times, compared with urban commuters.

These issues are of interest not only for train-operating companies but also for policy-makers because rail fare complexity implies a trade-off between two societal objectives: efficiency and equity. While price discrimination allows for a more efficient allocation of demand to the different services, by adjusting prices to customers' willingness to pay, it also has a negative influence on customers' perceptions of price fairness (Wu *et al.* 2012). The search costs to understand the fare system may also be regarded as unfair. In fact, Bonsall *et al.* (2007) argue that complex fare systems in the transport sector tend to be regarded as unfair more often than in other sectors, because they assume that travellers need to make an effort when choosing departure time, ticket type, and companies, in order to avoid paying what is perceived as a "price penalty". There is also evidence that this effort is particularly impactful in the case of older people and those less familiar with online booking systems (SPATF 2011, BDRC Continental 2015), which raises additional equity issues. Even when consumers are fully informed, price differentiation in public transport may still be rejected on social and moral grounds (Weesie *et al.* 2009).

The present study aims at filling the gap in the literature regarding the relationships between fare structure complexity and rail demand, by analysing the impact of different fare complexity scenarios on the choices made by current users and non-users of the interurban rail network and on overall rail demand along one of the main interurban rail routes in the UK (London-Leeds).

3. Survey design and sample characteristics

The main component of the research was a survey to users and non-users of the rail network. The sample of users was recruited from the database of online ticket sales of a rail company (Virgin Trains East Coast) and included individuals who had made at least one rail trip on the

London-Leeds route in the preceding month. The survey consisted of 1,027 online interviews, completed between 29 July and 3 September 2016. Non-users were interviewed using laptops in hall tests, with recruitment outside the venue, in two cities along the London-Leeds route (Leeds and Doncaster). Non-users were recruited on the basis of having made no trip by rail and at least one trip by car or coach longer than 50 miles along the Leeds-London route in the preceding month. A £5 voucher was provided to participants as an incentive. The survey consisted of 179 interviews, completed between 11 and 16 August 2016. Although the non-users sample is relatively small, the results are still fairly robust, as shown in the sections that follow. However, the small sample does limit our ability to capture the diversity of non-users.

The survey questionnaire contained questions about the demographic characteristics of participants, recent rail and non-rail trips on the London-Leeds route, general perceptions about the rail fare structure and the different ticket types, and a stated preference exercise where participants chose among ticket types and train services. Two pilot surveys were conducted, the first one including 81 users and 38 non-users and the second including 98 users. Minor changes were then made to the survey questionnaire applied in the main stage of the research.

Table 1 shows the characteristics of the samples of users and non-users, compared with the respective populations. The population values were derived from the UK National Travel Survey trips dataset (DfT 2017) and were calculated based on a definition of users and non-users similar to the one described above and aggregating individuals making trips of more than 50 miles from/to the regions along the London-Leeds rail route. The table shows that the composition of the two samples is roughly similar to the respective populations and reproduces the main differences between rail users and non-users that occur in those populations: on average, users are more affluent and make more trips, which tend to be longer and more frequently for commuting and business. The only difference between samples and populations is an imbalance between the proportions of the youngest (16-29) and oldest (60+) age groups in the users and non-users samples, compared with the populations.

Table 1: Characteristics of sample and population (%)

		Sample		Population	
		Users	**Non-users**	**Users**	**Non-users**
Gender	Male	57	60	61	62
	Female	43	40	39	38
Age	16-29	12	33	23	15
	30-44	23	28	67	66
	45-59	39	22		
	60+	26	17	10	19
Personal income (per year)	Under £20,000	16	60		
	£20,000 to £29,999	15	16		
	£30,000 to £49,999	29	14		
	£50,000 to £74,999	18	6		
	£75,000 or over	22	4		
	Under £25,000			42	58
	£25,000 to 49,999			32	31
	£50,000 or over			26	11
Number of trips last month on London-Leeds route (all modes)	1-2	60	74	55	75
	3-4	23	12	14	12
	5+	17	13	30	13
Purpose of most recent trip	Commuting	28	12	30	13
	Business	19	10	22	17
	Personal	53	78	49	69
Duration of most recent trip	0-39 minutes	1	5	0	0
	40-79 minutes	17	30	15	21
	80-119 minutes	45	30	39	38
	120 minutes or over	37	36	46	41

N: 1,027 users, 179 non-users

4. Perceptions about the rail fare structure and ticket types

Participants were then asked whether they agreed or disagreed with a series of statements regarding the rail fare structure (Figure 1). The majority of users agreed that the fare structure is poorly explained. Non-users had more negative perceptions, with the majority agreeing not only that the fare structure is poorly explained, but also that working out which is the best ticket is too complicated, and that there usually is a better value ticket for their trip that they are not aware of. Despite having these views, only a minority of participants (11% of users and 34% of non-users) agreed that the fare structure is so complex that it puts them off travelling by train, and only 28% of users and 42% of non-users would rather have a simplified system but excluding some low cost options.

Figure 1: Opinions about current rail fare structure
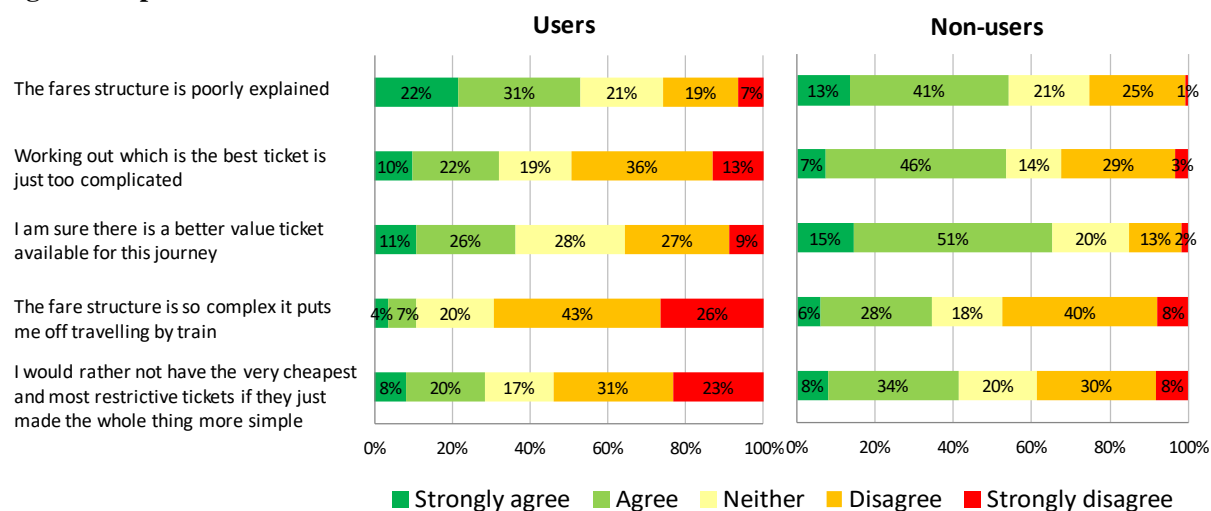
Participants were also asked about their awareness of different ticket types. They were first asked to think about a "base trip" by rail (for users, this was their most recent trip, and for non-users, a hypothetical future trip) and to indicate which ticket type they bought or would buy, and what other alternative ticket types they could have bought/could buy. 22% of users and 21% of non-users who bought/would buy Advance tickets for their base trip and 23% of users and 38% of non-users who bought/would buy Anytime tickets did not know what alternative ticket types they could buy.

Rail users were also asked about the most important considerations for the choice of ticket type. The main considerations were value and ticket cost, mentioned by 40% and 38% of users respectively. A smaller proportion (14%) mentioned flexibility.

There were some variations within the samples, as participants in the oldest age group (≥60 years old), in the lowest income groups (<£30,000/year), and making only 1-2 trips per month, and personal (rather than commuting or business) base trips consistently agreed to all the statements in Figure 1 to a larger extent than other groups. The groups with the worst perceptions about the fare structure were also the ones with lowest awareness of ticket types, with the exception of the oldest participants, who had better awareness of ticket types than those in the youngest group (<30 years old). Participants in the youngest age group, in the lowest income groups, and making personal trips also gave more consideration to buying the cheapest ticket than other groups, while individuals in the highest income groups (≥£50,000/year) and making commuting or business trips give more consideration to flexibility. There were few differences according to sex and duration of the base trip.

Overall, the results point to a negative general perception of the fare system and to a lack of awareness of how the system works, especially among people who currently do not travel by
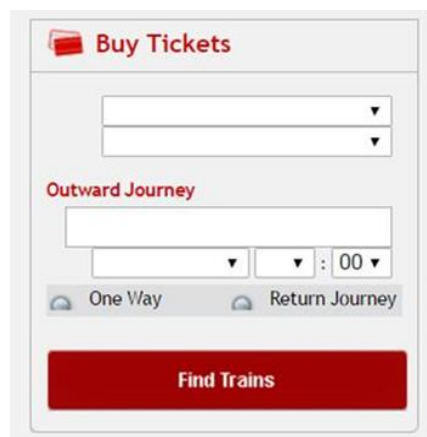
rail (non-users). However, this varies according to personal and trip characteristics. The analysis of stated preference data in the following sections investigate whether these perceptions lead to any significant impacts on demand for rail.

5. Stated Preference: design

5.1. Choice scenario

The main part of the survey was a choice experiment that mirrored very closely the actual booking experience faced by customers buying tickets online, using as a template the booking engine of one of the train companies. Participants were first shown a screen (Figure 2) where they selected the origin and destination of their trip and the ideal departure or arrival time and then clicked on the "return journey" button to add information about the return leg, if applicable. They were then presented with 8 choice situations (Figure 3).

Figure 2: Selection of trip details

Figure 3: Example of choice situation

In each choice situation, participants could choose among up to 531 different ticket type combinations for outward and return legs. These combinations were derived from 58 different individual ticket types offered on the route, including Advance, Anytime, Off-Peak, and Super Off-Peak tickets, valid for single or return trips, First or Standard class, and for travel on specific companies or on any company. The set of options that each participant saw in the experiment included all ticket types that are currently on offer on the route segment of the selected trip (except season tickets).

As well as choosing the preferred ticket type, participants could also choose from up to 25 possible train services for their trip, including a service at the stated ideal departure or arrival time and up to 12 earlier and later services (with time increments based on the frequency of real-world train services for the route segment in question). Participants who chose an Off-Peak or Super Off-Peak ticket or a ticket that was specific to a single company only saw train service options that were available on that ticket type.

The choice process replicated the real-world booking process. Participants saw a maximum of 10 ticket types and their fares for the outward leg of the trip and could click on "more fares" to see others (if available). On selecting one of the ticket types, the list of services that could be chosen with that ticket was displayed, together with text (below the date) describing the ticket restrictions. Participants then chose one of those services. If the ticket type for the

outward leg was a single ticket, then they could choose the ticket type and service for the return leg, using the same method. If the ticket type for the outward leg was a return ticket, the return leg window showed only train services, not ticket types. If the trip was a single trip, the return leg window did not appear. The selection was finally confirmed by clicking "Buy now".

Importantly, and in the one key departure from the real-world booking engine, participants could also indicate that they would not buy any ticket. This option allowed us to explore the impact of differing levels of complexity on the decision whether or not to travel by train.

5.2. Attributes and levels

The key design variable was a 'complexity' attribute, which took four levels, as shown in Table 2. These levels included the existing fare structure (Level 0), where the price of each type of Advance ticket varies from service to service, two simplifications to that structure (Levels -2 and -1), and one increase in complexity (Level +1). The levels match our definition of complexity, made in the introduction, as the range of different ticket types, with different prices and restrictions. *Complexity Level -2* (No Advance tickets) reduces complexity both by reducing the number of different tickets types available and removing the tickets with most restrictions (i.e. Advance tickets). *Complexity Level -1* (Same price for all the services available with each Advance ticket) reduces complexity by reducing the number of different prices. *Complexity Level +1* (new Advance Flexible tickets are available) increases complexity by increasing the number of different ticket types available.

Table 2: Complexity levels

Level	Name	Description
-2	No Advance tickets	Advance tickets are not available.
-1	All Advance tickets same price	The price of each Advance ticket type is the same for all train services that can be used with that ticket.
0	As Now	The price of each Advance ticket type varies from service to service.
+1	Flexible Advance tickets available	Flexible Advance tickets (First and Standard class) are added to the offer. These tickets are valid on the chosen service or on the service immediately before or afterwards.

The other attributes of the design were the fare and duration of the train services, both with 9 levels, from -20% to +20%, in 5% increments, relative to the real-world fare and duration of the ticket type that participants stated they bought/would buy for their base trip and the service with the preferred departure or arrival time for that trip. Within each complexity level,

the fares of Advance tickets varied by service, except at *Complexity Level -1*, where the Advance tickets offered by each company had a fixed fare (equal to the average of the fare of all the services that can be used with that ticket in *Complexity Level 0*). The fares for Flexible Advance tickets in *Complexity Level +1* were specified with an increment of 40% above ordinary Advance tickets for the same service. The design was also restricted so that Flexible Advance tickets were always cheaper than Anytime tickets. The fares of tickets other than Advance and Advance Flexible were the same in all complexity levels and did not vary by service.

5.3. Experimental design

It is considered best practice to derive the sequence of attribute levels presented to participants in each choice situation by using an efficient design, i.e. one that maximises the statistical precision of the model estimates. In the present case, however, the complexity of the design rendered this unfeasible using available software. This is because the choice structure had two levels, with the upper level comprising the ticket type alternatives and the lower level comprising the train services available with the chosen ticket type. In addition, the two choice sets depended on the complexity attribute, which constrained the alternatives shown in different ways depending on its level. At *Complexity Level +1* the design was unconstrained; at *Complexity Level 0* the design was constrained only insofar as none of the new Flexible Advance tickets was made available; at *Complexity Level -1* all Advance tickets were additionally constrained to have the same price across all services, and at *Complexity Level -2* all Advance tickets were removed from the set of available alternatives.

   We therefore used a combination of efficient design and random allocation techniques. The lower level exercise (choice of train service conditional on ticket type) was developed as a 'D-efficient' design (Rose and Bliemer 2014), using the *Ngene* software package. The priors that entered the algorithm for this component were drawn from the pilot surveys mentioned in Section 3. The upper level exercise was developed as a random design by randomly varying the fare levels of each ticket type around their base level within the range +/- 20%. Although the resultant overall design cannot lay claim to the desirable efficiency properties obtained under the D-efficient design method, it can still be expected to result in unbiased estimates and levels of precision that are sufficient for the purpose of the study. In fact, findings from the literature suggest that the experimental design technique is less crucial for model precision than the sample size and the set of attributes and levels (Louviere *et al.* 2013, p.20).

Considering the substantial sample size obtained, the experimental design approach adopted was judged to be the most appropriate solution for the present study. No significant correlations between attribute levels were found in the design.

6. Stated preference: analysis

6.1. Frequency of choices

Table 3 shows the frequency with which the "not train" alternative was chosen in the stated preference exercise, in each level of complexity. For users making return trips, that alternative was chosen in 14% of all choice situations with *Complexity Level 0* and *Complexity Level +1*. The proportion grew to 22% in *Complexity Level -1* and to 59% in *Complexity Level -2*. For non-users making return trips, the "not train" alternative was chosen in 42% of all choice situations with *Complexity Level 0*. The proportion was lower (32%) in *Complexity Level +1*, slightly higher (46%) in *Complexity Level -1* and much higher (86%) in *Complexity Level -2*. For both users and non-users, the frequency of choices for the "not train" alternative was always smaller in the case of single trips, in each complexity level. The relatively low proportion of non-users choosing the "non train" alternative for single trips may be explained by the overrepresentation of younger people in the sample, as they tend to have a higher predisposition to travel by train than other age groups (see for example DfT 2018).

Table 3: Frequencies of choice of the "not train" alternative, by complexity level (%)

Complexity level	Users		Non-users	
	Return trips	**Single trips**	**Return trips**	**Single trips**
-2 (No Advance tickets)	59	42	86	26
-1 (All Advance tickets same price)	22	18	46	11
0 (As Now)	14	9	42	4
+1 (Flexible-Advance tickets available)	14	11	32	6

Overall, the results suggest that the simplifications to the fare regime that were offered in the survey would have a negative impact on demand for rail travel. In particular, withdrawal of Advance tickets from sale altogether would be expected to have a substantial downward impact on demand. The models in the following sections investigate whether these results still apply when controlling for the other factors (such as the fare and the characteristics of the available ticket types) that may affect the choice for rail. The models also test whether increasing complexity by introducing Flexible Advance tickets has a positive or negative effect on train travel overall, as this is not clear from the results in the table above.

6.2. Model structure

The stated preference choice data was analysed using a nested model (Figure 4). The upper-level model explains the choices for ticket types and the lower level model explains the choices for train services conditional on ticket type choice. The ticket types have different types of restrictions regarding departure time (no restriction, Off-Peak time only, or Super Off-Peak time only) and whether they can be used on services offered by any train-operating company (TOC) or only on services by a specified company. Participants choosing a time-restricted ticket were only shown train services departing at the specified time. Participants choosing a company-restricted ticket were only shown services operated by that company.

Figure 4: Ticket type and train service choice models
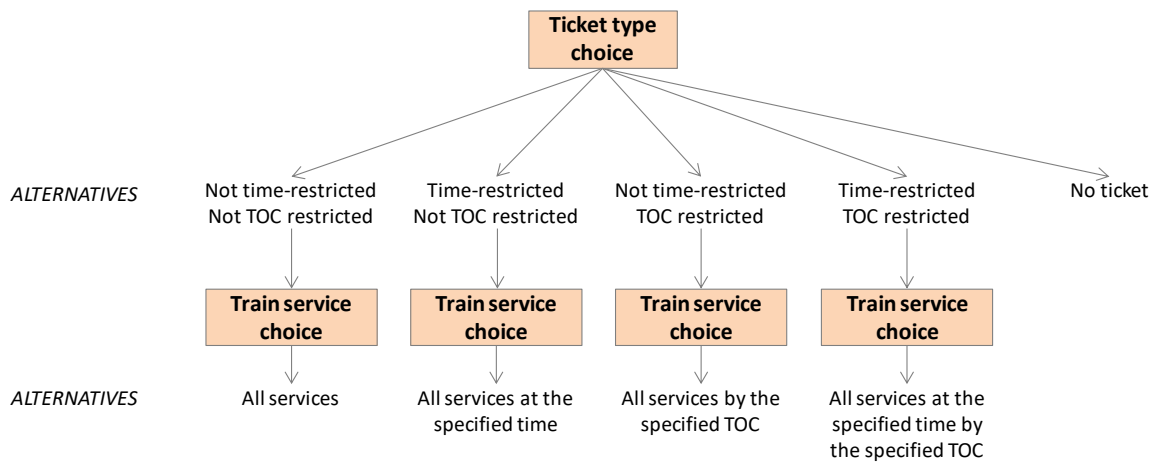
The analysis of the two choices (ticket type and train service) is linked by allowing the utility of the set of train services available with a given ticket type to enter into the ticket type choice model. This utility depends on the restrictions of each ticket type regarding travel on off-peak or super off-peak times and the company that can be used. This is explained in detail in the next section.

6.3. Model specification

Lower-level model (train service choice)

The lower-level model explains the choice of train service, and was estimated separately for users and non-users. Each observation represents one alternative train service for each trip leg in each choice situation faced by each participant. The number of alternative services available for each leg depends on the chosen ticket type.

The model was formulated as a mixed logit. The utility of a train service j to individual i was specified as

$$U_{i,j} = \beta_i \, x_j + \varepsilon_{i,j} \tag{1}$$

where x_j is a vector with the characteristics of train service j, β_i is a vector of individual-specific parameters to estimate, and $\varepsilon_{i,j}$ is an error term representing the idiosyncratic preferences of individual i towards train service j.

The probability that individual i chooses train service j and not any alternative service k in his choice set can be expressed as the probability that the utility associated with service j is greater than the utility associated with service k i.e.

$$P(U_{i,j} > U_{i,k}) = P(\beta_i \, x_j + \varepsilon_{i,j} > \beta_i \, x_k + \varepsilon_{i,k}) \tag{2}$$

If the error term follows a Type I Extreme value distribution, then the probability that individual i chooses service j, given all other services k, can be expressed in terms of a logistic distribution as follows:

$$P_{i,j} = exp(\beta_i \, x_j)/(exp(\beta_i \, x_j) + \Sigma_k exp(\beta_i \, x_k)) \tag{3}$$

This model can be estimated using maximum likelihood methods (Train 2009).

Table 4 lists the x_j variables entered in the model. We assume that the utility of a train service depends (negatively) on the fare, duration, and positive or negative differences between the departure or arrival time and the ideal departure or arrival time. We used a spline specification to account for non-linear effects of differences from the ideal departure/arrival time. This was done by adding variables accounting for positive and negative differences above 2 hours. All variables were specified with normally distributed random parameters, except the fare, which had a fixed parameter. Standard errors were corrected by grouping the observations belonging to the same participant.

Table 4: Variables in the train service choice model

Variable	Description
fare	Fare for the train service, conditional on the chosen ticket type, measured in £
duration	Duration of the trip, measured in hours
difference (pos)	Number of hours after the ideal departure or arrival time
difference (pos)>2	Number of hours in excess of 2 hours after the ideal departure or arrival time
difference (neg)	Number of hours before the ideal departure or arrival time
difference (neg)>2	Number of hours in excess of 2 hours before the ideal departure or arrival time

Upper-level models (ticket type choice)

The upper-level model explains the choice of ticket type and was estimated separately for users and non-users and single and return trips. The return trip models treat the choice of

ticket type for the outward and return legs as a single combined choice. Each observation represents an alternative ticket type for each choice situation faced by each participant. The number of ticket types available in each situation is different in each route segment, varying from 3 to 21 in the case of single trips and from 12 to 531 in the case of return trips (including return tickets as well as combinations of single tickets).

The explanatory variables are listed in Table 5. We assume that the utility of a ticket type depends negatively on the fare (or, in the case of Advance and Advance Flexible tickets, on the average fare across all services). The *cheapest* dummy variable allows for an "excess demand" for the cheapest ticket, i.e. for the probability of a participant choosing the cheapest ticket in any choice situation being higher than the probability predicted only by absolute fare levels. The *fare(cheapest)* variable accounts for interactions between that "excess demand" effect and the level of the cheapest fare.

Table 5: Variables in the ticket type choice model

Variable	Description
fare	Fare for the ticket type, measured in £
cheapest	Dummy variable indicating that a ticket type is the cheapest in the choice situation
fare (cheapest)	Fare for the cheapest ticket type, measured in £
logsum (out) and *logsum (return)*	Maximum utility the participant would expect to gain from its choice of train service in the outward/return leg if the ticket type in question was chosen
anyTOC (out) and *anyTOC (return)*	Dummy variables indicating that the ticket can be used on any train operating company (TOC) in the outward/return leg, rather than being specific to a single company
not train	Dummy variable indicating the "would not travel by train" alternative
not train[complexity=-2]*	Interaction between not train and complexity=-2, a dummy indicating that no Advance tickets were available
not train[complexity=-1]*	Interaction between not train and complexity=-1, a dummy indicating that the price of each Advance ticket type is the same for all train services that can be used with that ticket
not train[complexity=+1]*	Interaction between not train and complexity=+1, a dummy indicating that new Flexible-Advance tickets were available

We also assume that the utility of a ticket type depends positively on the maximum expected utility that can be derived from the set of train services that can be used with that ticket in the outward and return legs of the trip. This utility is assessed by a term called the 'logsum' (De Jong *et al.* 2007), which can be estimated from the train service model as the natural logarithm of the denominator of the logistic function in expression (3). For example, the logsum of a ticket type *t* for individual *i* in a given trip leg is given by

$$logsum_{i,t} = ln(exp(\beta_i x_j) + \Sigma_k exp(\beta_i x_k)) \tag{4}$$

where $\beta_i x_j$ and $\beta_i x_k$ are the utilities of services j and k. Tickets with more restrictions regarding departure time and train companies have lower logsums as the utilities of the restricted services are not entered in the calculation in expression (4).

Two dummy variables (*anyTOC out* and *anyTOC return)* identify ticket types that can be used in the services of all companies serving the route segment in question. This is to account for preferences for the services of specific companies, after controlling for the lower price and the loss in flexibility of company-specific tickets (which is accounted for by the logsum variables). There are no *a priori* expectations regarding the sign of the *anyTOC* variables.

The key variables in the model are the interactions between a dummy for the "not train" alternative and dummies for complexity levels -2, -1, and +1. The coefficients of the interactions measure the impact of those complexity levels on choices for the non-train alternative, in comparison with Level 0, and controlling for other factors affecting choices, such as fare, company, and, via the logsum terms, how the set of services that can be used with each ticket meets preferences regarding trip duration and preferred departure or arrival time. There are no *a priori* expectations regarding the coefficients of the interactions. As noted in the introduction and literature review, complexity can have a positive or negative impact on rail demand.

The choice is again modelled as a mixed logit, with a structure analogous to the one in expressions (1) to (3). The *fare* and *fare(cheapest)* variables and the three interactions between the *not train* variable and the complexity levels were entered with fixed parameters. All other variables entered the model with normally distributed random parameters.

6.4. Results: Train service choice models

Table 6 shows the estimated train service models for users and non-users. The *fare* coefficients are negative, as expected. The *duration* coefficients are also negative, as expected, although in the non-users model the coefficient is not statistically significant. The coefficients on *difference(pos)* and *difference(neg)* are negative, which indicates that participants prefer to depart/arrive close to their ideal departure/arrival time, again as expected. The *difference(pos)>2* and *difference(neg)>2* coefficients are negative, which shows that the preference for departing/arriving close to the ideal departure/arrival time, in comparison with departing/arriving earlier or later, is stronger when the difference is above 2 hours. The coefficients of the standard deviations are almost all significant, showing that there are variations in preferences within the sample.

Table 6: Lower level model (Train service choice given ticket type choice)

	Users		Non-users	
	Coefficient	*Std. Error*	*Coefficient*	*Std. Error*
Mean				
fare	-0.114	0.003***	-0.073	0.010***
duration	-0.428	0.063***	-0.027	0.175
difference(pos)	-0.898	0.050***	-1.990	0.149***
difference(pos)>2	-2.349	0.100***	-0.384	0.184**
difference(neg)	-1.926	0.069***	-0.960	0.171***
difference(neg)>2	-1.496	0.128***	-2.049	0.255***
Std. Dev				
duration	0.669	0.127***	0.334	0.416
difference(pos)	1.650	0.047***	2.129	0.138***
difference(pos)>2	2.412	0.084***	3.872	0.256***
difference(neg)	2.258	0.072***	3.236	0.189***
difference(neg)>2	2.056	0.104***	3.232	0.395***
Observations	129,072		25,151	
Choice situations	10, 528		1,727	
Participants	987		173	
Pseudo R^2	0.481		0.413	

Significance levels: ***: 1%, **: 5%, *: 10%

6.5. Results: Ticket type choice models

Table 7 shows the estimated ticket type choice models for users and non-users making return and single trips. The *fare* coefficients are negative, as expected. The coefficients of the *cheapest* dummy variable are positive (and significant in all models except the one for single trips of users). This shows that there is an "excess demand" for the cheapest tickets available in a choice situation, compared with the demand predicted by fare levels alone. The coefficient of the *fare(cheapest)* variable is significant and negative in the models of users and non-users making return trips. This shows that for these trips, the excess demand effect for the cheapest ticket type was weaker when the price of this ticket type was higher. The coefficient of the *fare(cheapest)* variable is significant and positive in the model for single trips of users. This shows that for these trips, the excess demand effect for the cheapest ticket type was stronger when the price of this ticket type was higher.

Table 7: Upper level model (Ticket type choice)

	Users				Non-users			
	Return trips		Single trips		Return trips		Single trips	
	Coef.	*Std. Err.*	*Coef.*	*Std. Err.*	*Coef.*	*Std. Err.*	*Coef.*	*Std. Err.*
Mean								
fare	-0.032	0.001***	-0.035	0.001***	-0.052	0.004***	-0.050	0.006***
cheapest	2.461	0.283***	0.053	0.186	3.039	0.420***	1.332	0.480***
fare (cheapest)	-0.030	0.004***	0.010	0.003***	-0.016	0.007**	-0.014	0.010
logsum (out)	0.230	0.013***	0.131	0.023***	0.098	0.045**	0.085	0.162
logsum (return)	0.099	0.043**			0.009	0.039		
anyTOC (out)	-1.449	0.067***	-1.222	0.089***	-1.813	0.265***	-1.124	0.298***
anyTOC (return)	-2.291	0.117***			-3.218	0.476***		
not train	-6.374	0.538***	-7.141	0.542***	1.863	0.863**	-7.192	1.374***
not train[complexity=-2]*	4.082	0.399***	3.309	0.377***	0.265	0.763	2.096	1.222*
not train[complexity=-1]*	1.113	0.220***	1.635	0.269***	-0.067	0.373	1.672	1.090
not train[complexity=+1]*	-0.040	0.218	0.403	0.270	-0.886	0.352**	0.465	1.051
Std. Dev								
cheapest fare	1.774	0.137***	2.149	0.145***	0.643	0.397	2.844	0.599***
logsum (out)	0.118	0.021***	0.330	0.041***	0.304	0.063***	0.868	0.142***
logsum (return)	0.625	0.040***						
anyTOC (out)	0.246	0.103**	0.454	0.141***	0.502	0.389	0.762	0.370**
anyTOC (return)	1.691	0.151***			2.046	0.406***		
not train	11.17	0.512***	6.416	0.416***	6.539	0.792***	6.209	1.354***
Observations	823,391		31,840		127,966		3,200	
Choice situations	4,527		3,034		889		362	
Participants	646		381		133		46	
Pseudo R²	0.433		0.380		0.695		0.467	

Significance levels: ***: 1%, **: 5%, *: 10%

The *logsum* coefficients are positive (although not all are significant). This shows that, as expected, participants prefer tickets with fewer restrictions, all else equal. Furthermore, the coefficients are less than one, which is consistent with economic theory (Ortúzar and Willumsen 2011, Ch.7.4.3).

The *anyTOC* coefficients are negative. This shows that participants prefer tickets that could only be used on services operating by a specific company, rather than those that can be used on any company, after controlling for the lower price and the loss of utility resulting from the restriction in the number of services available to choose from on these tickets. This result may account for a "status quo bias", as suggested by previous studies (Paha *et al.* 2013), or it may reflect strong preferences for the services provided by specific companies, due to factors such as reliability, comfort, and on-board amenities.

The coefficient of the *not train* variable is negative for users and for non-users making single trips. This shows that these participants prefer to make the trip by train, rather than using other means of transport. Non-users making return trips prefer not to make the trip by train.

The coefficient of the interaction between *not train* and *Complexity Level -2* is positive in all models (although not significant in the model for non-users making return trips). This shows that, all else equal, participants were more likely to choose to travel by train when Advance tickets were available than when they were withdrawn from sale.

The coefficient of the interaction between *not train* and *Complexity Level -1* is positive and significant in the models for rail users. This shows that individuals who currently use the rail network were more likely to choose to travel by train when the price of each Advance ticket type varied by the train services that could be used with that ticket, compared with a scenario when the price did not vary.

The coefficient of the interaction between *not train* and *Complexity Level +1* is negative and significant in the model for non-users making return trips. This shows that for these participants, the preference for travelling by train is stronger when there are additional Flexible-Advance tickets available.

The coefficients of the standard deviations are almost all significant, showing that there are within-sample variations in preferences for ticket types. To investigate the source of these variations, we estimated additional models, including interactions with dummy variables representing personal and trip characteristics (using the same classification as in Table 1). Table 8 shows the significant interactions between those characteristics, the 'not train' variable, and the different complexity levels. The results are consistent with the segmented analysis of perceptions about fare complexity reported in Section 4. Although there are differences in the set of significant variables across the four models, in general, groups that are more averse to fare complexity (e.g. the youngest and oldest age groups and individuals making personal trips) have higher propensity to choose rail in less complex scenarios (as shown by the negative coefficients in Level -2 and -1) and lower propensity in more complex scenarios (as shown by positive coefficients in Level 1). Groups that are less averse to complexity (e.g. those in the highest income groups) have lower propensity to choose rail in less complex scenarios (as shown by positive coefficients in Level -2). In addition, we found that interactions with trip duration were significant in three of the models, with individuals making longer trips showing higher propensity to choose rail in less complex scenarios (as shown by negative coefficients in Levels -2 and -1).

Table 8: Significant interactions between 'not train', complexity, and personal/trip characteristics

Complexity level	Users		Non-users	
	Return trips	**Single trips**	**Return trips**	**Single trips**
Level -2	Income 50-75k (+)	Age<30 (-)	Duration 80-120 mins. (-)	Duration 80-120 mins. (-)
	Income>75k (+)		Duration > 120 mins. (-)	Duration > 120 mins. (-)
	Personal trip (-)			
	Duration > 120 mins. (-)			
Level-1	Income>75k (+)	Commuting (-)	Age>60 (-)	
			Duration > 120 mins. (-)	
Level 1			Age>60 (+)	Age<30 (+)

Note: (+) positive coefficient; (-) negative coefficient. Significance level: 10% or less

7. Prediction of demand impacts

The models were then used to predict the rail demand associated with different fare structures (i.e. different levels for the complexity attribute) in five key segments in the London-Leeds route, which represent different trip durations and different numbers of companies and ticket types (Table 9).

Table 9: Route segment characteristics

Route segment	Average trip duration (minutes)	Number of different TOCs	Number of different ticket types available
London-Leeds	132	2	15
London-Doncaster	100	4	28
Leeds-Peterborough	86	2	12
Leeds-Doncaster	41	1	8
Peterborough-Doncaster	55	2	10

Note: route segments include both directions

The model was first adjusted so that the predictions in the "as now" scenario in each route segment were consistent with real-world data. The "real" shares of rail in the total number of trips in each of the five route segments were derived from the National Travel Survey trips dataset (DfT 2017), considering trips between the regions that contain the origins and destinations of the five route segments. The share of each ticket type in all rail trips was then estimated from the Latest Earnings Networked Nationally Overnight (LENNON) database, which contains information about the number of tickets sold, and their fares, for all train-operating companies in the UK, disaggregated by origin, destination, and ticket type.

We then used a calibration process which estimates alternative-specific constants for each ticket type in the ticket type choice models (following the approach in Train 2009, Ch.2.8). The utility of ticket type t for individual i in market m (i.e. users and non-users making single and return trips) in a given route segment was defined as

$$U_{i,t,m} = \alpha_t + \beta_{i,m} x_t + \varepsilon_{i,t,m} \tag{5}$$

where α_t is a constant measuring the characteristics of ticket type t that are not accounted for by the x_t explanatory variables (those in Table 5).

We then set an iterative process where at each stage v the constant was equal to

$$\alpha_t^v = \alpha_t^{v-1} + ln(S_t^*/S_t) \tag{6}$$

where S_t^* and S_t are the real and predicted preference shares of ticket type t. The constants are initially set to zero, i.e. $\alpha_t^0=0$. At each stage, the predicted share of ticket type t combined the shares in the four markets as below

$$S_t = S_r^* \ (p_{u,s} \ S_{t,u,s} + (1-p_{u,s}) \ S_{t,u,r}) + (1-S_r^*)(p_{nu,s}S_{t,nu,s} + (1-p_{nu,s})S_{t,nu,r}) \tag{7}$$

where S_r^* is the real market share of rail, $p_{u,s}$ and $p_{nu,s}$ are the proportions of single trips in the users and non-users sample, and $S_{t,u,s}, S_{t,u,r}, S_{t,nu,s}$ and $S_{t,nu,r}$ are the predicted shares of ticket type t in the four markets (i.e. users making single and return trips and non-users making single and return trips). These four predicted shares are calculated as the average of the probabilities that individuals in the sample chose ticket type t. For example, the share of ticket type t for users making single trips is

$$S_{t,u,s} = [\Sigma_i \ exp(\beta_{i,u,s} \ x_t)/(exp(\beta_{i,u,s} \ x_t + \Sigma_z exp(\beta_{i,u,s} \ x_z))]/n \tag{8}$$

where n is the sample size and x_z are the values of the explanatory variables in the utility functions of other ticket types z. The predicted shares of each ticket type eventually converge to the real shares ($S_t = S_t^*$), which means that the alternative-specific constants also converge to a final value ($\alpha_t^v = \alpha_t^{v-1}$).

In the calculation of the predicted shares for each route segment using expression (8), the value of the $\beta_{i,u,s}$ coefficients are the averages of the mixed logit individual-level coefficients of all participants i whose base trip was on that segment. The values of the explanatory variables x_t were assigned as follows. The fares are the real-world fares. The *cheapest* and *fare(cheapest)* variables were calculated after comparing the fares of all ticket types available in the route segment. The values of the *anyTOC* variables were assigned based on the characteristics of the ticket. The *logsum* values were calculated using expression (4), including, for each ticket type, the utility of the available train services in the route segment in question. As the calibration was run for *Complexity Level 0* ("as now"), the three interactions between the *not train* variable and the complexity dummy variables were all set to 0.

The calibrated models, with the final alternative-specific constants, were then used to predict the preference shares of the non-rail alternative and of each ticket type, in each of the five segments, for each complexity scenario.

Figure 5 shows the impact of complexity levels -2, -1, and 1 on rail demand and revenue in the five route segments, compared with *Complexity Level 0*. It also shows the impact of a scenario where *Complexity Level -2* occurs simultaneously with a reduction in all non-Advance fares so that the average fare for all available ticket types remains the same. This is to isolate the impacts of fare complexity and of the associated changes in fare levels, as Advance fares are in many cases the cheapest fares available. The main bars in Figure 5 show the percentage change in the overall demand for rail (i.e. the demand for all ticket types). The dotted lines show the impact on revenue.

If Advance tickets were not available (*Complexity Level -2*), the demand for rail travel would fall substantially, although the magnitude of the decrease varies by segment, from 11% to 45%. Even in the case when the removal of all Advance fares occurs simultaneously with a reduction in the remaining fares, there is still a reduction in demand of between 3% to 37%, which suggests that removing Advance fares has a negative impact that is independent of the fact that some of the cheapest fares are removed. If Advance tickets had the same price across all services (*Complexity Level -1*), the impact on overall rail demand would still be negative in all route segments, but likely to be small (3% to 6%). The introduction of new "Flexible Advance" tickets (*Complexity Level +1*) would increase the demand for rail travel in all route segments, with the magnitude of the increase varying from 4% to 15%.

Figure 5: Change in rail demand and revenue, by route segment and complexity level
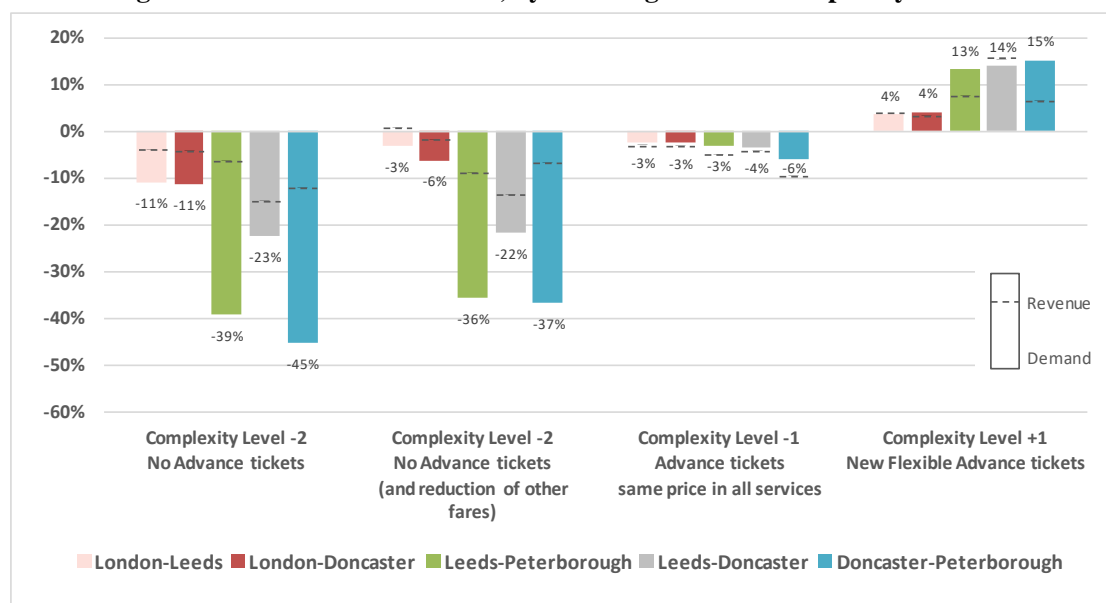
Note: Values were calculated as 100*(*Rail share/Rail share in Complexity Level 0 – 1*)

The impacts on revenue diverge from the impacts on demand due to differences in the prices paid by four groups of customers: those who stopped using rail, those who started using rail, those who continued using rail but chose a different ticket type, and those who continued using rail and the same ticket type. In *Complexity Level -2*, revenue decreases less than demand, in relative terms, because the customers who switched from Advance tickets to non-rail represented a less than proportional part of the revenue, since Advance tickets are generally cheaper than the average of all tickets. In addition, customers who switched from Advance Fares to other types of tickets generate more revenue than in the base scenario, as those other ticket types are more expensive. In the scenario where *Complexity Level -2* is accompanied by a reduction in other fares, we found that there was also a general switch towards more expensive tickets, which become more attractive due to the price reduction.

In Complexity Level -1, revenue decreases about as much as demand, because Advance tickets are priced at their mean level so some customers pay more for Advance fares and others pay less, compared with *Complexity Level 0*. In addition, we found that the customers who switched from rail to non-rail used to pay a price similar, on average, to the price paid by customers who did not switch.

In *Complexity Level 1*, customers who switched from non-rail to the new Advance Flexible tickets pay less than the average price paid by original customers, as Advance Flexible tickets are cheaper than the average of all tickets. This should imply a percentage increase in revenue smaller than the increase in demand. However, we found that in three of the segments, this was compensated by a switch to Advance Flexible tickets from cheaper tickets.

These predictions should be considered alongside some caveats regarding the extent to which the model captures preferences for using rail. The uncalibrated predictions of the share of rail in the "as now" scenario are close to the real shares in the route segments where this share is high (i.e. the two London segments) but more than 20% higher in the other segments. This may be explained by hypothetical or policy bias in the answers to the stated preference experiment (also evident in the non-users' low frequencies for the non-rail alternative in Table 3) or by measurement errors in the real shares (which are based on population data that is aggregated by region). Despite this caveat, the implied conditional fare elasticities of the overall rail demand derived from the models (-0.77 to -0.32, depending on the route segment) are broadly consistent with previous literature (Wardman 2014), so the model does not overestimate **changes** in the propensity for choosing rail. It is also likely that the demand impacts vary with the individual and trip characteristics shown in Table 8, but it was not

possible to calibrate the predicted shares by sample segment to the respective real-world shares because the population data (LENNON data) gives information only on the tickets sold and not on the characteristics of the buyer.

8. Discussion and conclusions

This paper has sought to enhance understanding of the demand impact of fare complexity in the interurban rail sector, starting from the hypotheses that a complex fare structure offers a greater opportunity to find the most appropriate ticket but it may also frustrate potential passengers and possibly inhibit demand. We used a stated preference experiment to analyse the impact of different complexity scenarios, defined by the number of different ticket types with different prices and restrictions, focusing on Advance tickets and using the London-Leeds route as case study. Nested mixed logit models were used to estimate the impact of complexity on the decision to travel by train, while controlling for fare levels, service durations, differences from ideal departure/arrival time, and restrictions on the services that can be used with each ticket. The models were then used to predict changes in the overall rail demand attributable to the different fare regimes examined, after calibration to real-world data in five key route segments.

We found that simplifying the fare structure by removing Advance tickets would have a substantial negative impact on demand (11% to 45%, depending on the segment), and a smaller impact on revenue, even when this occurs with a simultaneous reduction in all remaining fares. Equalising the fares of all Advance tickets on a given route segment would have a small negative impact on demand (3 to 6%), driven mainly from the behaviour of rail users (i.e. survey participants who had made a rail trip in the preceding month). In contrast, increasing complexity by introducing Advance Flexible tickets would have a positive impact (4-15%), in this case driven mainly from non-users making return trips. It is also likely that these results vary with personal characteristics (especially age and income) and trip characteristics (especially purpose and duration).

The results confirmed the hypothesis that customers like the flexibility provided by having many options to choose from, which offsets the negative aspects that have been identified in the literature, such as search costs and the risk of choosing the wrong ticket. More complex fare scenarios, with more ticket types, increase the chances that individuals find their preferred ticket type, i.e. their preferred combination of price and restrictions. This happens as the range of ticket types is successively augmented with Advanced tickets (i.e. going from

Complexity Level -2 to -1), then with Advanced tickets with different prices (Level -1 to 0), and finally with Advance Flexible tickets (Level 0 to 1).

However, we have also found that the higher propensity for choosing rail when the choice scenarios are more complex occurs alongside a general negative perception about the complexity of the fare system, as shown in Section 4. This could be because the choice scenarios are about particular trips, defined by the participants, who may already have some knowledge about the available ticket types and the "best deal", while general perceptions are influenced by more general factors, such as demographics, political context and the influence of the media. Negative perceptions of complexity may also depend less on the number of different ticket types than on the lack of clear information about the restrictions for those tickets. For example, we found that the majority of both users and non-users agreed that the current fare structure is poorly explained.

The paper thus adds to the literature on consumer behaviour by confirming that people prefer to have a wide range of alternatives to choose from, but that they also need these alternatives to be clear, in order to reduce the costs and risks of the choice. This supports giving a higher policy priority to providing better information than to reducing the number of different ticket types, a conclusion which is consistent with the current policy view in the UK (see DfT 2013, p.12 and ORR 2015). Better information could be provided by personalised online booking systems, using data on past behaviour, and more straightforward and interactive interfaces (e.g. voice recognition). The findings are also relevant beyond the UK, especially in Europe, where many companies already sell Advance tickets, and where the variety of tickets may increase following the upcoming market deregulation. The results of the complexity scenario involving Advance Flexible tickets are particularly relevant in countries such as Ireland, where there are different types of Advance tickets with different degrees of flexibility, and also apply in other sectors where these tickets are already well-established, such as air travel. More broadly, our conclusions are relevant in sectors, such as hotel bookings or mobile phone contracts, that increasingly depend on online sales and where customers are typically faced with a large number of options.

The paper also makes a methodological contribution to the study of complex choices made online, by using a stated preference experiment that replicated the situations that individuals face in real-world booking engines. While this method implies showing participants repeated choice situations with many alternatives, there was no evidence of respondent fatigue (for example, incomplete questionnaires or non-trading behaviour). Although it is still difficult to

derive efficient designs for this type of experiments involving many alternatives, the modelling of the choices produced robust estimates. The model results also give some hints regarding the heuristics that individuals use to make decisions, for example, choosing the cheapest tickets or tickets valid for a specific company. The methods could be used to test other changes to the fare structure (such as removing Off-Peak tickets or return tickets) or to estimate the demand for different services throughout the day. The approach could also be expanded to analyse choices over more than one trip (with season tickets as options), and over how far in advance tickets are bought, as well as the effect of different charging mechanisms, including simplifications (e.g. distance or zone charging) or increases in complexity (e.g. dynamic pricing).

Acknowledgements

[removed in submitted version]

References

Abrate, G., Piacenza, M., Vannoni, D. (2009) The impact of integrated tariff systems on public transport demand: evidence from Italy. *Regional Science and Urban Economics* **39 (2)**, 120-127.

Ariely, D. (2000) Controlling the information flow: effects on consumers' decision making and preferences. *Journal of Consumer Research* **27 (2)**, 233-248.

BDRC Continental (2015) Rail Ticket Retailing: The Passenger Perspective. Report to the Office of Rail Regulation., http://orr.gov.uk/__data/assets/pdf_file/0007/18187/Rail-ticket-retailing-passenger-perspective-report170615.pdf

Bonsall, P., Shires, J., Maule, J., Matthews, B., Beale, J. (2007) Responses to complex pricing signals: theory, evidence and implications for road pricing. *Transportation Research Part A: Policy and Practice* **41 (7)**, 672-683.

Booz & Co. (2009) *The Benefits of Simplified and Integrated Ticketing in Public Transport*. Report to the Passenger Transport Executive Group., http://www.pteg.net/system/files/general-docs/integratedticketingreportFINALOct09.pdf

Castillo-Manzano, J I., Sánchez-Braza, A. (2011) An evaluation of the establishment of a taxi flat rate from city to airport: the case of Seville. *Urban Studies* **48 (9)**, 1909-1924.

Dargay, J M., Pekkarinen, S. (1997) Public transport pricing policy: empirical evidence of regional bus card systems in Finland. *Transportation Research Record* **1604**, 146-152.

DfT (UK Department for Transport) (2013) Rail Fares and Ticketing: Next Steps., https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/249001/fares-ticketing-next-steps.pdf

DfT (UK Department for Transport) (2017) National Travel Survey 2002-2016 (Data Collection). 12th Edition. UK Data Service. SN: 5340., http://doi.org/10.5255/UKDA-SN-5340-8

DfT (UK Departmemnt for Transport) (2018) National Travel Survey: England 2017., https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/729521/national-travel-survey-2017.pdf

De Jong, G., Daly, A., Pieters, M., Van Der Hoorn, T. (2007) The logsum as an evaluation measure: review of the literature and new results. *Transportation Research Part A: Policy and Practice* **41 (9)**, 874-889.

Ellsberg, D. (1961) Risk, ambiguity and the savage axioms. *Quarterly Journal of Economics* **75**, 643-669.

ESA Retail (2018) Research into Passengers' Experience of Ticket Vending Machines. Report to Office of Rail and Road., http://orr.gov.uk/__data/assets/pdf_file/0016/24046/research-into-passengers-experience-of-ticket-vending-machines-february-2017.pdf

Garbarino, E C., Edell, J A. (1997) Cognitive effort, affect and choice. *Journal of Consumer Research* **24 (2)**, 147-158.

Hess, S., Shires, J., Bonsall, P. (2013) A latent class approach to dealing with respondent uncertainty in a stated choice survey for fare simplification on bus journeys. *Transportmetrica A: Transport Science* **9 (6)**, 473-493.

Iyengar, S S., Lepper, M R. (2000) When choice is demotivating: can one desire too much of a good thing? *Journal of Personality and Social Psychology* **79 (6)**, 995-1006.

Kahn, B E., Lehmann, D R. (1991) Modeling choice among assortments. *Journal of Retailing* **67 (3)**, 274-299.

KPMG (2018) Towards a Future Fares Strategy. Report to the Rail Delivery Group., https://www.raildeliverygroup.com/files/Publications/2018-05_towards_a_future_fares_strategy.pdf

Lee, B-K., Lee, W-N. (2004) The effect of information overload on consumer choice quality in an on-line environment. *Psychology and Marketing* **21 (3)**, 159-163.

London Economics (2015) Ticket Retailing Code of Practice – What Information is Relevant. Report to the Office of Rail Regulation., http://orr.gov.uk/__data/assets/pdf_file/0008/14795/lse-ticket-retailing-code-of-practice-report-july-2014.pdf

Louviere, J J., Carson, R T., Burgess, L., Street, D., Marley, A A J. (2013) Sequential preference questions factors influencing completion rates and response times using online panel. *Journal of Choice Modelling* **8**, 19-31.

Matas, A. (2004) Demand and revenue implications of an integrated public transport policy: the case of Madrid. *Transport Reviews* **24 (2)**, 195-217.

ORR (Office of Rail and Road) (2015) A Code of Practice on Retail Information for Rail Tickets and Services., http://www.nationalrail.co.uk/static/documents/content/RetailInformationForRailTicketsCoP.pdf

ORR (Office of Rail Regulation) (2012) Fares and Ticketing – Information and Complexity., http://orr.gov.uk/__data/assets/pdf_file/0018/4905/ticket-complexity-report-june-2012.pdf

Ortúzar, J D., Willumsen, L G. (2011) *Modelling Transport* (4th Edition). Wiley, Chichester.

Paha, J., Rompf, D., Warnecke, C. (2013) Customer choice patterns in passenger rail competition. *Transportation Research Part A: Policy and Practice* **50**, 209-227.

Passenger Focus (2010) Ticket Vending Machine Usability - Qualitative Research., http://www.transportfocus.org.uk/research/publications/ticket-vending-machine-usability-qualitative-research

Passenger Focus (2012) Passenger Focus Response to the Government's Rail Fares and Ticketing Review., http://www.transportfocus.org.uk/research/publications/passenger-focus-response-to-the-governments-rail-fares-and-ticketing-review

Preston, J., Whelan, G., Wardman, M. (1999) An analysis of the potential for on-track competition in the British passenger rail industry. *Journal of Transport Economics and Politics* **33 (1)**, 77-94.

Preston, J. (2008) Competition in transit markets. *Research in Transportation Economics* **23 (1)**, 75-84.

Reibstein, D. J., Youngblood, S. A., Fromkin, H. L. (1975) Number of choices and perceived decision freedom as a determinant of satisfaction and consumer behavior. *Journal of Applied Psychology* **60 (4)**, 434-437.

Rose, J M., Bliemer, M C J. (2014) Stated choice experimental design theory: the who, the what and the why., in S Hess and A Daly (Eds.) *Handbook of Choice Modelling*. Edward Elgar, Cheltenham.

SDG (Steer Davies Gleave) (2016) Study on the Prices and Quality of Rail Passenger Services. Report to the European Commission Directorate General for Mobility and Transport., https://ec.europa.eu/transport/sites/transport/files/modes/rail/studies/doc/2016-04-price-quality-rail-pax-services-final-report.pdf

Sharaby, N., Shiftan, Y. (2012) The impact of fare integration on travel behavior and transit ridership. *Transport Policy* **21**, 63-70.

SPAFT (SPA Future Thinking) (2011) Understanding and Testing Passenger Perceptions of Complexity in Relation to Fares and Ticketing. Report to the Office of Rail Regulation., http://docplayer.net/5619612-Understanding-and-testing-passenger-perceptions-of-complexity-in-relation-to-fares-and-ticketing.html

Swait, J., Adamowicz, W. (2001) The influence of task-complexity on consumer choice: a latent class model of decision strategy switching. *Journal of Consumer Research* **28 (1)**, 135-148.

Taylor, S R H., Carter, D W. (1998) Maryland Mass Transit Administration fare simplification: effects on ridership and revenue. *Transportation Research Record* **1618**, 125-130.

Train, K. (2009) *Discrete Choice Methods with Simulation* (2nd Edition), Cambridge University Press, Cambridge.

Wardman, M. (2014) Price elasticities of surface travel demand - a meta-analysis of UK evidence. *Journal of Transport Economics and Policy* **48 (3)**, 367-384.

Wardman, M., Toner, J. (2003) Econometric modelling of competition between train ticket types. *White Rose Consortium ePrints Repository.*, http://eprints.whiterose.ac.uk/2439/1/ITS2211-Econometric_modelling_upload.pdf

Weesie, L., Hofker, F., Kroes, M., Mitrani, A. (2009) The impact of tariff differentiation on time of day choice and rail demand in the Netherlands. *Association for European Transport Papers Repository.*, https://aetransport.org/public/downloads/IlltT/3850-514ec5c6bef5f.pdf

Whelan, G., Batley, R., Shires, J. (2008) Optimal fares regulation for Britain's railways. *Transportation Research Part E – Logistics and Transportation Review* **44 (5)**, 807-819.

Which? (2011a) Why Train Fares Don't Add Up - Which? Finds Train Staff Give Poor Ticket Advice., http://www.which.co.uk/news/2011/02/why-train-fares-dont-add-up-245320

Which? (2011b) Many people don't understand train tickets; rail ticket websites add to confusion., http://www.which.co.uk/news/2011/09/many-people-dont-understand-train-tickets-265963

Wu, C-C., Liu, Y-F., Chen, Y-J., Wang, C-J. (2012) Consumer responses to price discrimination: discriminating bases, inequality status, and information disclosure timing influences. *Journal of Business Research* **65 (1)**, 106-112.

Yang, C-W., Chang, C-C. (2011) Applying price and time differentiation to modeling cabin choice in high-speed rail. *Transportation Research Part E – Logistics and Transportation Review* **47 (1)**, 73-84.