Dissanayake D.

[Watching the clock on the way to work? Analysing trends in commuting activities, modes and gender differences in commute times, using hazard based duration modelling methods.](#)

Journal of Transport Geography 2017, 65, 188-199.

Copyright:

© 2017. This manuscript version is made available under the [CC-BY-NC-ND 4.0 license](#)

DOI link to article:

<https://doi.org/10.1016/j.jtrangeo.2017.10.013>

Date deposited:

21/11/2017

Embargo release date:

03 May 2019

This work is licensed under a

[Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence](#)

Watching the Clock on the Way to Work?

Analysing trends in commuting activities, modes and gender differences in commute times, using hazard-based duration modelling methods

---

Abstract

This study uses Hazard-based duration modelling methods to investigate commute patterns of males and females, with the aim of exploring the variability of commute times and modes over a period of nine years beginning in 2003. Securing two major datasets of the UK National Travel Survey (NTS) and Tyne and Wear household travel survey, as well as the level of detail that the chosen datasets offer, made it possible to ascertain the complexity of commuter travel at its fundamental level. Duration models have been very popular when analysing duration related activities. However, duration modelling research in the context of transport so far has been restricted to cross sectional one-off datasets. This is the first study that investigates commuting durations at a disaggregate level over a sustained period of nine years using duration modelling methods to acquire a fundamental understanding of the changes in commute durations. Gender aspects as well as the transport modes, including non-motorised transport (NMT), car and public transport (PT), were also addressed in the analysis. Probability density functions (PDFs), survival functions (S(f)) and hazard functions (H(f)) were employed when carrying out in-depth investigations into the patterns of commuting activities generated by males and females for the whole period of study over nine years, followed by year on year analysis. The descriptive analysis shows that commuting times are becoming longer as time goes by. Year on year analysis reveals that commute trips by car made by males are more likely to prevail in the system compared to PT, especially towards the end of the study period in 2010-11. The opposite is true for female commuting trips. NMT was an attractive mode for both males and females in 2009-10 even for longer commute trip durations of over 50 minutes. As the complexity of the activity travel patterns of males and females has not been given adequate attention in previous research, this study made a step forward in investigating the gender aspects, with specific attention being given to the differences in commute times. Transport authorities' and policy makers' timely interventions, such as quality bus partnerships, cycle city guides, and cycle to work schemes, as well as Metro system reinvigorations, including smart ticketing, have been taken into account when rationalizing the results and the changes to commute times over the study period.

Keywords: Commuter travel; males and females; duration models; travel behaviour.

---

1. Introduction

The changing nature of commute behaviour has received considerable attention in recent years from researchers, planners and policy makers alike due to its potential association with social and environmental concerns. Lawton (1), in one of the earliest investigations in this area, identified the growth in commute distances in Great Britain as being problematic. Since then, researchers have continued to investigate the changes in commuter travel. However, over the past several decades, increased attention has been paid to this subject due to increasing vehicle ownership rates and its adverse impacts on our society, economy and the environment, leading to problems such as urban sprawl, traffic congestion, travel delays, traffic related emissions and health issues. Case studies conducted in France (2), Germany (3), Italy (4), Netherlands (5), Spain (6), Sweden (7) and USA (8, 9) have explored the connection between urban structure and commuting patterns. Among recent studies in the UK, investigations into car ownership, commute travel mode, and time (10), the changing nature of commute patterns (11) and commuting trends (12) have contributed considerable knowledge to the subject of commuter travel.

The amount of time that commuters spend on their daily commute is vital to the evolution of urban systems in general, and to transport and land use patterns in particular. Easy access to the work place and the duration of the trip to work are fundamental to commuters in relation to decisions concerning where to live and where to work (13). Long commute journeys, especially by car, have been found to generate an adverse influence on health and

well-being (13, 14). Longer commutes lead to lower levels of productivity at work, as well as financial burdens due to increased transportation costs (13). In addition, negative utility generated by travel has been a major topic of discussion in transport research since the 1970s (15). Therefore, travel patterns and their associated travel time warrant careful attention for comprehensive travel demand modelling and forecasting procedures to aid strategic decisions; for instance, value of travel time savings, transport system improvements, policy interventions, economic appraisal and transport investments. Commuting trips account for a greater proportion of the overall demand for travel in any country and therefore this study focuses on commute time to explore how it changes over an extended period of time.

According to the study by Dainton (11), the percentage of commuter travel in the UK that takes less than 20 minutes has fallen since 2003. As reported by the UK Department for Transport (16), the average commute time was 27 minutes in 2006; that was an increase of 13% from the 1995/1997 figure of 24 minutes. It was reported that the number of workers who commute daily for two hours or more in the UK increased by a third during 2011-16, with stagnant wages and soaring housing prices pushing people further away from their work places (17). It is reasonable to assume that increased congestion on the network during peak periods may also be a possible reason for longer commute times. A study conducted in Minnesota (US) examined the reasons for large increases in commute times, with an 11% increase in the 1990s over the 1980s average figures. It ruled out land use factors and travel speeds from the set of factors influencing the change, but pointed to unemployment rates and gasoline prices in the 1990s as possible reasons for the changes in commute times (18). According to the European Working Conditions Survey (19), time spent on commuting journeys over complete home-work-home cycles in Europe has been in the range of 31.6 - 53.9 minutes. Within this range, the UK was ranked to be the 6th highest, with an average commute time of 45.4 minutes, out of 31 countries included in the survey.

The concept of “gendered mobility” has been popular in the context of transportation research in recent years. The differences between male and female commuters regarding their travel behaviour with respect to social interactions, work related commitments, household responsibilities and activity participation have been given attention in order to answer questions on how much, by what means and for what purposes they travel (20-22). Early studies in the UK have identified that there are some similarities and differences between males and females with regard to their commute behaviour (11, 23). As reported recently, female commuters in the UK have shorter commute times than males (11). Past research confirms that males spend more time travelling than females (24-27). There was a consistent trend observed for females in the UK until the mid-1990s regarding their tendency to opt for walking or using slower forms of public transport such as buses and trams (23). Past research provides useful information about the differences between genders in commute behaviour. However, most of them investigate the issues in a broader sense rather than providing a comprehensive explanation of changes in commute behaviour of males and females over time.

Since there is currently a limited understanding of the changing nature of commute times over time, further research is required. This study therefore seeks to address the issue through analysis of household travel data collected by UK National Travel Survey (NTS) coupled with a more detailed household travel survey in the Tyne and Wear Region in North East England during 2003-11. This will help to investigate how commute times have evolved over the years and how they can be linked into current and future policies to help promote a more sustainable urban environment. Hazard based duration technique is considered to be a suitable modelling tool in this analysis, given the nature of the analysis that is required. Previous research findings confirm that there is a substantial difference between males and females in how they organize their commuting journeys. Therefore the models were estimated separately to investigate gender differences in making commuting trips in general, and commute times in particular. The study uses the commute trips between home and work (and vice versa) without intermediate stops. The dataset does not include details of trip chaining behaviour and, due to the data limitations, the study focuses on direct home-work and work-home trips only.

This paper begins by presenting a critical review of previous research covering long-term changes in commuting time and modelling approaches for analysing commuting time data. The theory and suitability of Hazard based duration methods in transport research is also discussed. Attention is then directed towards a discussion of the study area and data description. This is followed by a presentation of the methodological approach and results of the analysis. The paper concludes with an explanation of potential reasons for changing commute times over time with attention to the case study area and directions for future research.

2. Analysing long-term changes in commuting time: A review of previous studies

The evolution of commuting time over a long-term period has been a point of discussion in previous research as a basis for policy making decisions. However, the definition of “long-term” has been rather ambiguous. Many studies making reference to “long-term analysis” only selected 2-3 discrete time points, typically at the beginning, middle and end of the overall study period (28-31). By doing so, they may have overlooked significant events in the interim periods due to environmental, political or economic changes. At least one study highlighted the benefits of performing analysis over a shorter time scale rather than the usual decennial census periods (28). Therefore analysis targeted at more frequent intervals, for example annually, is considered as having greater value.

In general, the changes in commute times have been characterized only by descriptive statistics, such as means and standard deviations, rather than considering more detailed analytical techniques, thereby taking advantage of the available datasets at disaggregate level (28-29). This may be due to the fact that the objectives of those studies were somewhat diverse. For example, Levingston and Wu (29) considered travel times for all trip purposes where a detailed analysis of commute times may not have been the main focus of their research. Similarly, the study by Saadi et al. (28) represented the changes to commute times by using spatio-temporal analysis with emphasis on data visualization. Therefore detailed analysis of changes to commute times was not presented in their research.

Analysis of commute durations using disaggregate data (e.g. individual level) is considered to be more useful as it has better predictive capabilities when compared to aggregate level analysis. However, only a few studies in the context of commute time have carried out disaggregate analysis (for example 8-9, 31). Ordinary Least Square (OLS) regression (8-9, 30) or logistic regression (31) were seen to be the most popular methods of analysis in previous research when exploring the changes in commute times in selected cross sections. However, researchers have highlighted some drawbacks in using OLS and logistic (or logit) methods when analysing duration data (32-33). OLS methods assume that the residuals (or error terms) are normally distributed; however, this assumption is considered to be unrealistic in the context of duration data where data often exhibit asymmetry, particularly if some observations have very long durations (right skewed) (33). In order to resolve this issue some researchers transformed the dependent variable by taking the natural log before applying OLS methods (31). This remedy mitigates the skewness issue, but it will not overcome other serious problems (32-33), such as:

- *Negative Predicted Values* (OLS may return negative predicted values even though this is unrealistic – durations must be positive)
- *Censoring Issue* (OLS does not easily distinguish between ‘censored’ and ‘uncensored’ observations. Dropping censored observations may create sample selection problems; this may also generate biases on more shorter durations in the datasets)
- *Time Varying Covariates* (OLS cannot easily accommodate independent variables or covariates that change value over time)

To overcome these issues, a modelling strategy should be specifically designed when analysing duration data. The distribution of event times, for instance the way in which commute times spread in datasets, is often far from normal, and is affected by the restrictions stated above. This would indicate that survival (duration) models are usually superior to OLS regression (32-33). Joly (34) further confirms that only duration analysis is designed to account for the dependence of the choice of whether to terminate travel on the duration of the endogenous variables themselves. The likelihood that a trip will be ended depends upon how long the trip has lasted. Duration models are particularly suitable to investigate the duration processes in which temporal dynamics need to be included (34).

3. Application of hazard based duration models in analysing commute times

The hazard based duration model has been popular in biostatistics and economics research since the 1980s. Hensher and Mannering (35) was one of the earliest applications that used hazard based duration models in the context of transport. The duration model focuses on an end-of-duration occurrence; for instance, end of commuting activity participation, given that the duration has lasted for a specified time (35). The hazard function can be used as a tool to compare different scenarios so that their differences can be investigated.

Hazard based duration models are developed considering fully parametric, semi-parametric and non-parametric methods. The choice between non-parametric, semi-parametric, and fully parametric methods for estimating duration models is seen as a complicated issue by researchers. According to Washington et al. (36), most of the transportation-related applications in this context have utilised either a parametric or semi-parametric hazard approach. Non-parametric applications are considered to be appropriate if there is little information available on underlying (or baseline) distribution; for example, lack of theory to decide the distribution or a small sample size (36).

Non-parametric models are generally rare in transportation research to date. An exception includes Bhat (37) which applied non-parametric hazard methods to investigate the durations of shopping activities in the Boston region of the United States. The results indicated that there is a significant interaction between the households' socio-demographic characteristics and changes in the generation of shopping activities. Komma and Sirinivasan (38) examined home-based commute timing decisions by flexible full-time workers using hazard models and data from a travel survey conducted in San Francisco in 2000. In this study, the baseline hazard model was developed using non-parametric distributions. The results indicate that travel duration has a significant effect on the choice of departure time. The semi-parametric model is considered as a useful method when there is only a little information available about the underlying hazard distribution. According to Washington et al. (36), it has two major drawbacks: firstly, the duration effects may not be recognizable and secondly, it may have potential loss of efficiency. Mohammadian and Doherty (39) applied a semi-parametric hazard model to predict the duration of time between planning and execution of pre-planned activities based on attributes of activity and characteristics of decision makers, considering Toronto as a case study.

In a fully parametric context, various forms of distributions should be considered before deciding upon the most suitable one. According to Washington et al. (36), the appropriate distribution should be selected based on theoretical grounds. Zhong and Hunt (40) investigated the Best-Fit Hazard Functions for Urban Weekend Activities, considering Calgary as a case study. They found that best-fit distributions for travel-related activity are lognormal or three-parameter lognormal types and for social and out-of-town activities they are Weibull or three-parameter Weibull types. Oh and Polak (41) used a Weibull distribution hazard model to study household activity duration using data from a detailed seven-day activity travel diary from the UK National Travel Survey. They used different types of activities; for example, work, shopping, personal business, social and recreation. Schonfelder and Axhausen (42) studied the rhythms of travel using hazard-based duration models where a six-week travel diary was administered in two German cities (Halle and Karlsruhe) in 1999 as part of the MobiDrive project. In this study, three families of hazard model were tested; namely fully parametric with Weibull specification, semi-parametric and non-parametric.

Previous research suggests that a fully parametric hazard model can be considered as a suitable model for activity and travel related modelling. However, the research conducted so far in this context has predominantly used the data collected at a defined time or at a particular cross section. Cross-sectional studies may not be able to provide the details about cause-and-effect relationships as they do not take into account what happens before or after the time considered. The benefit of longitudinal studies is that they will help detect the developments or changes in the characteristics of the commute times. In other words, longitudinal studies extend beyond a single moment in time. Therefore, they can be used to establish sequences of events. The unique datasets identified for the study facilitate the exploration of some differences in commuting times through a year on year analysis reported in this paper. It is expected that this study may serve as a starting point for future more detailed research on commuting behaviour in general, and commute times in particular.

4. Model specification – Hazard based duration models

Washington et al. (36) provides a detailed description of the model specification of hazard based models. The hazard function starts with the cumulative distribution function, $F(t)$:

$$F(t) = P(T < t) \quad (1)$$

where P is a probability, T is a random time variable and t is some specified time. In the case of daily activity choices by an individual, the cumulative distribution function is defined to indicate the probability of an individual planning a specific activity before some transpired time, t .

The first derivative of the cumulative distribution function with respect to time represents the corresponding density function and can be written as:

$$f(t) = \frac{dF(t)}{dt} \quad (2)$$

The hazard function $h(t)$, the conditional probability that an event will happen between t and $t+dt$, given that the event has not occurred up to time t , is given by

$$h(t) = \frac{f(t)}{1 - F(t)} \quad (3)$$

Another important function that is used to explain the analyses of hazard based duration models is the survivor function. It represents the probability that an individual remains in the state (survives) until time t .

$$S(t) = P(T \geq t) = 1 - F(t) \quad (4)$$

Consequently the hazard function is related to the survivor function as follows:

$$S(t) = \frac{f(t)}{h(t)} \quad (5)$$

The pattern of the hazard function has some important implications for the analysis. Considering the underlying event and the distribution of duration of that event, the hazard function may vary in shape, including monotonically increasing, monotonically decreasing, monotonically increasing first and then decreasing, and constant over time. In each case, different distributions can be chosen for the hazard functions. The distributions are generally in the form of Exponential, Log-logistic, Lognormal, Weibull, Gompertz, Inverse Gaussian, and Gamma types.

5. Case study area

5.1. The Tyne and Wear region in North East England

This study considers the Tyne and Wear region in North East England as a case study area. The Tyne and Wear (T&W) region is comprised of five districts: Newcastle, Gateshead, North Tyneside, South Tyneside and Sunderland (Figure 1). It has high-quality public transport alternatives such as bus, rail and Metro systems. The region has a population of 1,104,141, as recorded in 2011 (43), of which 55% are of working age (44).

Like most regions in the UK, one of the challenges in the Tyne and Wear region would be to reduce the level of traffic congestion and traffic related emissions. Car ownership in the region has remained lower than the national average. Hence, the potential for future car ownership growth in the region is high. It is now increasing at a rate twice the national average. As reported recently, one of the most congested roads outside of London in the UK is in the North East region; the peak hour speeds in North East are reported to vary from 16 to 28 mph during peak hours leaving commuters in stressful situation (45). Another recent article supports this notion by placing Newcastle and Sunderland cities at 10th position among the most congested cities in the UK in terms of commuter travel in 2016 (46). Despite of the increasing car ownership rates there, road capacity was considered as a main reason behind this congestion issue. The Highways England was set to deliver £15bn investment on improving roads, upgrading motorways and adding extra lanes to the road network to help improve congestion and journey times (45). As per Department for Environment Food and Rural Affairs (DEFRA) in the UK, all 'air quality zones' in the North East region are currently breaking legal air pollution limits (47). Possibility of implementing a pollution charging scheme is currently at the discussion stage (48). Under local Transport Plan (LTP) scheme, Newcastle city council has been active in monitoring air quality in the area; two air quality management areas are located in the areas so that the pollution levels can be monitored in regular basis. In addition Newcastle City Council encourages sustainable travel in the area, including walking, cycling and public transport. Also better and safer facilities, bus priority measures as well as upgrading traffic signals have been implemented to ensure better co-ordination and to promote alternative forms of travel to reduce congestion in the region (47).

The regional transport statistics published by the UK Department for Transport (DFT) were used in this study to gain a fundamental understanding of the changes of commute times in the Tyne and Wear region in North East England (TSGB, 49). As the duration of an activity may influence travellers' decisions regarding their selection of a mode of travel, commute times by travel mode for the period of 2004-13 in the Tyne and Wear region were examined (see figure 2).

According to figure 2, commute times change greatly according to travel mode. Rail has the highest commute time with an average of 40 minutes and walking has the lowest durations with an average of 12 minutes. When each mode is taken into account, the durations have been relatively stable over the time period 2003-07 as a whole. There was a slight decrease in commute times by car and rail during 2008-09 and 2012-13. In contrast, bus and walking have shown some increase in commute times during these two time periods. During 2008-09 and 2012-13, there was a sharp fall in real GDP in the UK economy (50). The reason for the changes to the commute times may be attributed to the reason that commuters moved from car to walking and from rail to bus to avoid commute expenses due to the great recession. Increase in commute times by bus may also be due to two main reasons.

Firstly, commuters spend more time on the roads during their daily commute because of rising traffic congestion. Secondly, commuters choose to live away from the urban areas to have a better quality of life and to gain benefits from lower house prices in suburbs. The possible reason for increasing commute times by non-motorised modes would be again due to traffic congestion on the roads; for example, commuters who used motorised modes may switch to non-motorised modes to avoid congestion on the roads even though they may need to walk for longer distances.

Fig. 2 Commute times by travel mode in the Tyne and Wear region (2004-13).
Produced using the data from TSGB (49)

5.2. Data used for the activity analysis

This study uses the data collected during 2003-2011 by Newcastle City Council (T&W Household Travel Survey – T&W HHTS). In 2003, the data was collected only from the North Tyneside, South Tyneside and Sunderland districts. This has been extended to cover all five districts in Tyne and Wear since 2004. The survey was conducted as a partial panel survey as some of the households participated in the survey throughout, but some did not. The databases consist of household, household member and trip information. The households' and individuals' demographic data comprise household size, number of infants, number of cars, motorcycles and bicycles owned, gender, age, driving licences held and employment status. The data indicates that in T&W the average household size is 2.13 persons, there is an average of 0.79 cars per household and 39.3% of households have no car available as of 2006. Three-quarters of respondents over 16 years hold some form of driving licence and 48% of adults are in employment.

The travel related activity data includes a range of trip information including origin, destination, mode of travel, journey purpose, start and end time and trip distance. According to the T&W HHTS Report (44), the data shows that around 85% of trips were intra-zonal trips. Home-based trips are the most frequently observed, accounting for 40% of all trips. Approximately 60-70% of trips are made by car, and the bus is the most frequently used mode of public transport, representing 15-20% of all trips. The morning period with peak occurrence of trips was between 7-9am and the evening period was between 5-6pm for all trips. 32% of all trips were less than 1km and the average distance of the journey-to-work trip was 7km. This is shorter than the 8.7 miles average trip length for the UK as reported by Dainton (11).

For the purpose of this study, home-to-work commute activities were extracted from the main database. Commute trips that contain complex trip chains have not been included in the analysis as this may divert the focus of the study. Also the respondents' main commuting journeys were considered in this analysis. On average, 700-900 commute trips per year have been randomly selected for the data modelling. This is because the number of commuting trips in each year was found to be different and therefore samples needed to be in a similar range to allow better comparison. There were fewer commute trips recorded during 2009-2011 period, when compared to 2003-2008. It is postulated that this was due to less funding contributed to data collection as the survey contract was reaching to its end.

Accordingly, the commuter database consists of 701, 768, 901, 836, 703, 718, 612, 490, and 446 commute trips for 2003-2011 respectively. Household car ownership is considerably high in the T&W region as 86% of household own one or more cars. Also cycle ownership is rising in the region and over the study period where 36% of the household own one or more cycles. Regarding the mode of travel for commuting trips, car/van shows a greater share (69.8%), followed by bus/coach (18.0%) and non-motorised modes (12.2%). As per Table 5, more than 50% of the commuting trips are within 0-5km range. This is comparatively shorter compared to the UK average figures. A recent report stated that shorter commuting distances are to be expected in many cities in the UK to keep up with region's industrial legacy for current economic patterns (51).

Table 1 Database characteristics

Variable	Shares (%)
Gender	
Female	48.1
Male	51.9
Age (years)	
16-24	13.3
25-34	20.7
35-44	25.7
45-54	25.2
55-59	13.1
60-64	1.8
65-69	0.2
Household car ownership	
0	15.2
1	43.1
2	33.9
3 or more	7.8
Household cycle ownership	
0	64.1
1	14.1
2	11.9
3 or more	9.9
Travel mode	
Bus/coach,	18.0
Car/van	69.8
Cycle	1.8
Walk	10.4
Distance travelled	
0-2km	23.0
3-5km	30.3
5-10km	25.8
10-20km	17.2
20-30km	2.5
30-50km	1.1

5.3. Comparison of UK National Statistics and the data used for the study

To investigate the suitability of the commuter travel dataset from T&W HHTS, an attempt was made to compare it with the data available via DFT national statistics (TSGB, 49). As the travel time statistics for the North East region were available in TSGB (49), it is a very useful and direct source for comparison here. Accordingly, the distribution of commute times (figure 3), and the differences of commute times by gender (figure 4) were investigated.

Figure 3 presents the distribution of commute times in the T&W region for the period 2004-11 with both datasets mentioned above. Assessing the comparability of the sample with the general populace, it is clear that the sample achieves a close fit. Over a third of commuters take 20 minutes or less and about two-thirds make it less than 40 minutes, while less than a fourth have commuting times greater than 40 minutes. The percentage of people commuting for less than 20 minutes had fallen during 2005-06 and 2009-10. The percentage of people commuting more than 41 minutes was higher in 2006, 2009 and 2011 and was lower during 2007-08, and 2010.

Fig. 3. Commute times in the Tyne and Wear region (2004-11).

When commuting travel by males and females are taken into account, females are more likely to make shorter commute times than males (figure 4). It is also observed that there was a gradual increase in commute times for both males and females over the period of 3 years starting from 2003. The commute times for males were fairly stable over 2008-2010 over 2008-10, whilst the same for females were somewhat unstable over the whole period of the study.

Fig. 4. Commute times by males and females in the Tyne and Wear region (2003-11).

A paired-samples t-test was conducted to compare the yearly mean commute durations of two datasets with respect to male and female commuters. The results indicate that there was no significant difference in the scores for the DFT National Dataset and the T&W HH Travel Survey Dataset regarding male and female commute durations. This indicates that the T&W HH Travel Survey is appropriate to be used in the analysis in this study. The relevant statistics are as follows:

- **Males:** The scores for the DFT National Dataset ($M=23.44$, $SD=2.24$) and the T&W HH Travel Survey Dataset ($M=22.67$, $SD=1.14$) conditions; $t(8) = 1.492$, $p = 0.174$.
- **Females:** The respective scores for the DFT National Dataset ($M=22.67$, $SD=0.87$) and the T&W HH Travel Survey Dataset ($M=22.37$, $SD=1.22$) conditions; $t(8) = 1.00$, $p = 0.347$.

6. Methodology

The methodology consists of three main stages: data preparation for the analysis, selection of best fit distribution, and data analysis with Hazard-Based Baseline Models. Figure 5 illustrates the activities involved in each stage and the sequence in which the activities are linked together to generate the intended outputs in each stage. The details of the stages are explained in the following sections.

6.1. Data preparation

As a first step, three data files from the household travel survey (household data, person data and trip data), were combined together in a single database. In so doing, relevant household, person and residential data can be referred together with the trip data as required in the analysis. Since this study focuses mainly on commuter travel, the commute trips (direct trips between home and work) were extracted from the main dataset. Analysing intermediate stop making behaviour of the commuters, for instance trip chaining behaviour of commuters, is not the purpose of this study and therefore such trips were not included in the analysis. Each commute trip contains the start and end times on a continuous time scale together with the origin and destination locations defined by the postcodes. The commute times were calculated by taking the difference between end and start times. Subsequently, the data was checked for consistency. The outliers were removed from the database. The data was coded as necessary for the analysis. The resulting database consists of 6,175 commute trips collected over a nine-year period since 2003.

Fig. 5 Methodology framework.

6.2. Selection of the best fit distribution

The shape of the hazard function is based on the distributional assumptions that are considered for the probability density function (52). The choice of the distributional alternative is decided upon theoretical grounds or statistical evaluation (Washington et al., 36). If the selected distribution is incorrect, the analysis will generate an inconsistent estimate of the baseline hazard (53). According to Washington et al. (36), this will affect not only the shape of baseline hazard, but also the efficiency and inaccuracies of the estimations. Therefore, it is vital to decide upon the best distribution to represent the duration process with respect to commute trips in this study. Since the purpose of the study is to investigate the dynamics of commute times over the years, the analysis is mainly focused on developing baseline hazard models; in other words, the distribution tests were entirely based on commute times with no attention being given to the external variables available in the database. Accordingly, the commute times for males and females for the period of 2003-11 were used to test the distributional alternatives. Among the options available such as Weibull, Lognormal, Exponential, Log-logistic, 3-Parameter Weibull, 3-Parameter Lognormal, 2-Parameter Exponential, 3-Parameter Log-logistic, Smallest Extreme Value, Normal, and Logistic; 3-parameter distributions have been excluded from the distribution fitting exercise as they have negative thresholds and are therefore inappropriate to use in the context of commuting durations. The best fit distribution was decided by considering two widely acceptable methods: adjusted Anderson-Darling test statistic and correlations coefficients. The main criterion for the selection of a distribution is with the lowest Anderson-Darling value and the highest correlations coefficient. The analysis indicates that the best is Weibull, followed by Lognormal, Log-logistic and Exponential distributions for the survey period (Figure 6). Therefore, Weibull distribution is selected for analysing data.

Fig. 6 Probability plots.

6.3. Data analysis using baseline hazard-based models

Using the hazard model with Weibull baseline specification, the commute times were analysed. The baseline hazard models are based on the assumption that all explanatory variables (covariates) are equal to zero. The Probability Density Function (PDF), Survival function ($S(f)$) and Hazard function ($h(f)$) were produced to investigate the patterns of commuting activities generated by males and females for the whole period of 2003-11, followed by year on year analysis. The commute modes, such as car, public transport (PT), and non-motorised transport (NMT), were also given due attention in the analysis.

The PDF presents the shape of the chosen distribution and informs the relative chance for a random variable, for instance a commute time, to occur at a given point. The $S(f)$ simply indicates the probability that the commute has not yet been completed by time t . In other words, $S(t)$ denotes probability of surviving (or remaining in the system) beyond time t . According to Washington et al. (36), the slope of the $h(f)$, in other words the first derivative with respect to time, is an important indication of the nature of the duration process under study. The $h(f)$ presents the probability (or likelihood) of ending a commute trip as a function of how long it has lasted. The hazard plot demonstrates the trend in the failure rate over time. Decreasing, constant and increasing hazard rates indicate that commute times are less likely to end as they increase in duration, are failing at a constant rate, and are more likely to fail as they increase in duration, respectively.

7. Results – Analysis based on the dataset covering 2003-11

7.1 Male and female commute times during 2003-11

The Probability Density Function (PDF), Survival function ($S(f)$) and Hazard function ($h(f)$) with respect to male and female commuters are shown in Figure 7. The PDF clearly indicates that females made shorter commute times compared to males in T&W Region in North East England during the entire period of the study. This result coincides with the previous findings (24-27, 49). The $h(f)$ s for both male and female commuters are non-monotonic. They are increasing up to 17 minutes for males and 13 minutes for females, and decreasing thereafter with time. It is also observed that the pace of increase of hazard rate is higher for females than males for short commute trips. This suggests that females are more likely to make shorter commute trips compared to males. On the other hand, the hazard rate appears to be higher for females for trips shorter than 38 minutes compared to males. This trend changes after 38 minutes, where the hazard rates seems to be somewhat higher for males compared to females. This indicates that the female commute trips longer than 38 minutes have a slightly better survival rate than men. In other words, the longer commute trips made by females are less likely to end as they increase in duration.

Fig. 7. Probability density, survival and baseline hazard functions for male and female commuting activities during 2003-11

7.2 Male and female commute times by transport mode during 2003-11

The PDF , $S(f)$ and $h(f)$ with respect to male and female commuters are shown in figures 8 and 9. The PDF describes the relative likelihood for the commute time variable to take on a given value. In other words, the values on the PDF therefore only provide an idea to inform about the spread. Whereas the $h(f)$ is very useful when compare different groups. For example, in a hazard model, the hazard rate of one group can be represented as a multiplier of the hazard rate of another group. In terms of the purpose prediction, $h(t)$ is considered to more useful compared to PDF .

Car was the most popular mode for males (75%) and females (64%) in T&W region regarding their commute trips followed by PT and NMT modes. Females use more PT and NMT alternatives compared to males. It is clear from PDF s in figures 8 and 9 that both males and females mostly use NMT for short trips, cars for medium range trips and PT for long trips. Regarding NMT, 90% of trips are between 0-35 minutes for males and 0-25 minutes for females respectively. 6 min (NMT), 10 min (Car) and 30 min (PT) min are the most frequent commuting trip durations made by females. Regarding males, the comparable figures to females are 10 min (NMT), 14 min (Car) and 34 min (PT).

According to figures 8 and 9, it is clear hazard functions for both male and female categories are increasing regardless of the mode of travel they use. However PT mode shows an increasing rate and the other two modes, NMT and car, have decreasing rates. This indicates that the trends of commute times by PT are different to NMT and car. In other words, commute trips by car and NMT are more likely to prevail in the system compared to the commute trips that are made by PT when the journeys are getting longer. To investigate further the reasons behind this outcome, year on year analysis was conducted in the next section.

Fig. 8. Probability density, survival and baseline hazard functions for male commuting activities during 2003-11

Fig. 9. Probability density, survival and baseline hazard functions for female commuting activities during 2003-11

8. Results – Analysis based on Year-on-Year basis for 2003-2011

The next step in the analysis was to investigate the changes in commute times for males and females on an annual basis. A year on year analysis is very useful as it allows for exploration of the changes in male and female commute travel patterns during the eight years of the survey: 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010 and 2011. However, the analysis is with small sample sizes when it comes to year by year basis. The sample size will be further reduced when the analysis is taken forward to investigate different modes on an annual basis. The other issues that might affect the analysis are the rounding effects when the respondents declare the commute times. Figure 10 shows that there is a tendency for the dataset to display rounding effects.

Fig. 10. Histogram to investigate the rounding effect of the dataset 2003-11

The only way to deal with this rounding effect and related issues is to consider large samples in the analysis. The confidence intervals for Weibull distribution is further studied to make sure the true effect has been captured from the analysis. Larger sample sizes would tend to generate more precise estimates; this would result in narrower confidence intervals. Figure 11 presents the probability plots for the dataset, both for 2003-11 and on an annual basis.

Fig. 11. Probability plots showing the Confidence Intervals

The Confidence Intervals are relatively narrow for both cases, indicating that the true effect of the data can be captured and the sample size is large enough to use for the year on year analysis (Table 2). However, a precautionary measure was taken forward in the analysis by combining the data for the years, for example 2003-05 (as year 2004), 2004-06 (as year 2005), so that the larger samples can be generated for the analysis. As the data collection occurred from April-March cycles for each year, this would not make any adverse impact on the analysis regarding how the datasets were combined to represent each year of the study period. The data integration has resulted in 2364 (2004), 2497 (2005), 2434 (2006), 2254 (2007), 2031 (2008), 1818 (2009) and 1547 (2010) samples for each year of the analysis.

Table 2 Characteristics of the dataset in terms of percentiles and Confidence intervals

Percent	2003-2011 dataset				Year 2003 as an example			
	Percentile	SE	95% CI (L)	95% CI (U)	Percentile	SE	95% CI (L)	95% CI (U)
1	1.672	0.070	1.539	1.815	1.200	0.157	0.928	1.552
5	4.470	0.132	4.218	4.737	3.438	0.317	2.868	4.120
10	6.902	0.167	6.582	7.238	5.472	0.415	4.715	6.349
50	21.514	0.270	20.992	22.048	18.467	0.731	17.088	19.957
90	44.398	0.493	43.443	45.374	40.092	1.402	37.436	42.937

8.1 Male commute times by transport mode – year on year analysis

The results are appended in figure 12. The *PDFs* show how male commuters decided upon modes depending on their commuting durations; for example, NMT for short, Car for medium, and PT for long durations. It can be observed that hazard functions are increasing, with a decreasing rate for car, and with an increasing rate for PT. The highest hazard rates for PT were observed for the years 2009 and 2010. This gives an interesting finding here in that the commuting trips by car are more likely to prevail in the system compared to the commute trips that are made by PT when the journeys are getting longer, particularly for the 2009-10 period. Does this raise the important question of whether the male commuters tend to use cars for commuting more than PT, especially for commuting longer durations to avoid higher PT fares during the time of recession in the UK?

Fig. 12. Probability density and baseline hazard functions for male commuting activities— year on year analysis

Regarding NMT, hazard functions increase with decreasing rate. It is also noted that lower hazard rates were observed for 2010 followed by 2009 regardless, of the commuting duration. Also, it was noted that hazard rates for NMT journeys (0.03~0.12) are comparatively lower than Car journeys (0.03~0.20) and PT journeys (0.03~0.24) for 2009-10. This is a particularly notable difference for longer journey durations. This observation adds to our understanding that the commute journeys by NMT became popular for longer journeys among male commuters in years 2009-10. This might be due to the travel awareness campaigns in T&W region to encourage travellers to walk and cycle. In addition, Cycle City Guides, with details including cycle maps were produced and advertised widely by Newcastle City Council in 2008 for the benefit of travellers who intend to use cycles for work, school and other purposes. On the other hand, the UK's great recession may have made NMT an attractive choice for male commuters.

8.2 Female commute times by transport mode – year on year analysis

The analysis is then focused on investigating the changes in commute times of females on an annual basis with respect to transport modes in the system. Figure 13 illustrates the results obtained from a year on year analysis over the study period.

Fig. 13. Probability density and baseline hazard functions for female commuting activities – year on year analysis

According to the *PDFs*, female commuters choose NMT for short, car for medium, and PT for long commuting durations in a similar way to male commuters. The hazard functions are increasing, with decreasing rates for car and NMT, and with an increasing rate for PT. Also, it is observed that the hazard rates for 2009-10 are comparatively lower than other years for PT, particularly for journeys of more than 50 mins as opposed to the trend observed for males as shown in figure 12. This also applies to the hazard rates for NMT for 2009-10, but this is true for all journeys regardless of the journey durations; this is comparable to the journeys made by males as shown in figure 12. It was observed that hazard rates for PT Journeys (0.03~0.13) is lower than that of car journeys (0.03~0.18). This gives an interesting finding here in that the commuting journeys by PT are more likely

to prevail in the system compared to the commute trips that are made by car for the 2009-10 period. This situation can be explained by the efforts of the local authorities in T&W region where some important changes were implemented on PT systems in and around 2008; for example, the T&W Metro reinvigoration, the launch of Quality bus partnership in T&W and the introduction of workplace travel plans.

When considering NMT journeys with PT and Car journeys, the hazard rates are considerably higher for longer commute durations. However, a significant reduction on hazard rate is observed in 2010 for NMT journeys compared to Car and PT journeys for longer commute durations (Figure 13). The introduction of cycle to work and cycle to school schemes under T&W Local Transport Plans (LTP), and the workplace travel plans may have contributed in a positive way to this change. Active participation of females in travel to work surveys indicates their engagement with community transport schemes (54). The ten year strategy 2011-22 for delivering cycling improvements in Newcastle has a main focus on those perceived to be under-represented in our cycling community, such as women, disabled people and ethnic minority groups, providing support and active encouragement for cycling (55).

9. Possible reasons for the changes in commute times and recommendations for future research

The descriptive analysis indicates that both males and females had shorter commute times in 2003 compared to other years; also, it is revealed that there is a higher chance that longer trips stay in the system in recent years; for example, in 2010 for male commuters (Figures 3 and 4). There may be three main reasons to support the findings:

1. The commuters relocate from the city to suburbs to avoid higher housing prices.
2. The traffic congestion in Tyne and Wear has followed an increasing trend since 2003.
3. Due to the great recession in the UK, car commuters may have changed to more sustainable options of NMT/PT modes.

One of the future aims of this research is to investigate the changes to the commute trip distances in T&W region in the study period. This will help to explore the reasons for the changes of commute times over the study period.

The 2nd local transport plan (LTP) in Tyne and Wear has made some efforts towards reducing traffic congestion in the region, aiming to deliver person journey time improvements. According to the LTP delivery report 2006-11, 10 transport corridors, out of 16 corridors selected under the LTP, have shown some significant reduction in journey times. However, this may have a counter-effect on the LTP's efforts on some additional measures such as encouraging modal shift to public transport alternatives including bus and metro systems. If car users changed to other modes in the system, including bus, metro, and slower but more sustainable modes such as bicycle and walking, the commute times may become longer. This will require a detailed investigation leading to an exploration of a potential connection between longer commute times in recent years and modal shift towards sustainable modes.

Regarding male commute journeys, men tend to use car over PT. In contrast, female commuters are more likely to use PT over car for commuter travel. The recent study about the PT services in London suggested that females are more prone to change their travel behaviour compared to males (56). NMT modes became more popular, even for longer journeys, for male and female commuters in 2009-10, indicating the effort that was made by the Local Authorities on promoting sustainable transport and smart cities initiatives. The possible reasons for this would also be due to heavy congestion on the roads, active travel initiatives by the local authorities, travel awareness campaigns, work place travel plans and support systems available from community groups (cycle hubs and route guides).

Recent improvements in the PT systems in Tyne and Wear region (for example, Superoute services provided via quality bus partnerships) may be a possible reason for attracting car travellers to PT alternatives. On the other hand, the recent initiative of metro reinvigoration and smart ticketing may also be leading to a modal shift towards the Metro alternative. The people who originally commuted by car for long distances may have switched to PT modes so that the commute times by PT may be longer than they were before. The UK's great recession since 2008 may have some significant influence on the changes in travel times in the cities and regions. However, can this be taken to have been a positive influence on the changes in the way we travel, with a possible modal shift of commuters towards PT and NMT modes? The further study will look in to investigating the reasons for the changes in commute times identified in this study.

References

1. Lawton, R. The Journey to Work in Britain: Some Trends & Problems. *Reg. Studies*, Vol 2, 1968, pp.27–40.
2. Aguilera, A. Growth in Commuting Distances in French Polycentric Metropolitan Areas: Paris, Lyon and Marseille. *Urban Studies*, Vol. 42, No. 9, 2005, pp. 1537–1547.
3. Scheiner, J. Interrelations Between Travel Mode Choice and trip Distance: trends in Germany 1976-2002. *Journal of Transport Geography*, Vol. 18, 2010, pp. 75-84.
4. Travisi, C. M., Camagni, R., and P. Nijkamp. Impacts of Urban Sprawl and Commuting: A Modelling Study for Italy. *Journal of Transport Geography*, Vol. 18, No. 3, 2009, pp. 382-392
5. Susilo, Y. and Maat, K. The Influence of Built Environment to the Trends in Commuting Journey in the Netherlands, *Transportation*, Vol. 34, 2007, pp. 589-609.
6. Gracia-Palomares, J. C. Urban Sprawl and Travel to Work: The Case of the Metropolitan Area of Madrid. *Journal of Transport Geography*, Vol. 18, 2010, pp. 197-213.
7. Sandow, E. Commuting Behaviour in Sparsely Populated Areas: Evidence from Northern Sweden. *Journal of Transport Geography*, Vol. 16, 2008, pp. 14-27.
8. Press, J. E. Spatial Mismatch or More of a Mishmash? Multiple Jeopardy and the Journey to Work. In *Prismatic Metropolis: Inequality in Los Angeles*, Russel Sage Foundation, 2002. Chapter 12, pp.453-488.
9. Shen, Q., Social and Spatial Dimensions of Commuting, *Journal of the American Planning Association*;66(1), 2000
10. Titheridge, H. and P. Hall. Changing Travel to Work Patterns in South East England, *Journal of Transport Geography*, Vol. 14, 2006, pp. 60-75
11. Dainton, E. The UK Commute: *Healthy or Hazardous? An Analysis of Current Commuting Trends*, National Travel Survey. <http://www.dft.gov.uk>. Accessed Nov. 2, 2010.
12. Dargay, J. and M. Hanly. Volatility of Car Ownership, Commuting Model and Time in the UK, *Transportation Research A*, Vol. 41, 2007, pp. 934-948.
13. Commute Time Trends in the CMAP Region, 2015. http://www.cmap.illinois.gov/about/updates/policy/-/asset_publisher/U9jFxa68cnNA/content/commute-time-trends-in-the-cmap-region
14. National Statistics, Does Commuting Affect Well-being? <https://www.ons.gov.uk/ons/>
15. Yamada, H. On the Theory of Residential Location: Accessibility, Space, Leisure, and Environmental Quality. *Papers of the Regional Science Association*, Vol. 29, 1972, pp.125-135
16. DfT. *Travel to Work*, National Statistics, Department for Transport, UK. 2007.
17. Trade Union Congress, Daily Commute of Two Hours – Reality for 3.7m UK Workers, The Guardian, 2016. www.theguardian.com/money/2016/nov/18/daily-commute-of-two-hours-reality-for-37m-uk-workers
18. Barns, G., Reasons for recent large increase in commute durations, Minnesota DOT, 2007.
19. European Working Conditions Survey, 2007. www.eurofound.europa.eu/pubdocs/2006/98/en/2/ef0698en.pdf.
20. Polk, M. *Gender Mobility, A study of Women's and Men's Relations to Automobility in Sweden*. Department for Interdisciplinary Studies, Human Ecology Section, Göteborg University, 1998.
21. Uteng, T. P. and T. Cresswell. *Gendered Mobilities*, Aldershot:Ashgate, 2008.
22. Frandberg, L. and B. Vilhelmson, More or Less Travel: Personal Mobility Trends in the Swedish population Focusing Gender and Cohort. *Journal of Transport Geography*, Vol. 19, 2011, pp. 1235-1244.
23. Pooley C. G. and J. Turnbull, Modal Choice and Modal Change: The Journey to Work in Britain Since 1980. *Journal of Transport Geography*, Vol. 8, 2000, pp. 11-24.
24. Prendergast, L. S. and R. D. Williams. Individual travel time budgets. *Transportation Research A*, Vol. 15, 1981, pp. 39-46.
25. Kitamura, R., J. Robinson, T. F. Golob, M. Bradley, J. Leonard and T. van der Hoorn. *A Comparative Analysis of Time Use Data in the Netherlands and California*. Research report number UCD-ITS-RR-92-9, Institute of Transportation Studies, University of California, Davis, 1992, June.
26. Levinson, D. and A. Kumar. Activity, Travel and the Allocation of Time. *APA Journal*, Autumn, 1995, pp. 458-470.
27. Robinson, J. *Time for Life. The Surprising Ways Americans Use Their Time*. The Pennsylvania State University Press, University Park, Pennsylvania, 1997.
28. Saad, I., Boussauw, K., Teller, J., and Cools M. Trends in Regional Jobs-housing Proximity Based on the Minimum Commute: The Case of Belgium, *Journal of Transport Geography*, 57, 2016, pp. 171-183.
29. Levinson, D. and Wu, Y. The Rational Locator Reexamined: Are Travel Times Still Stable? *Transportation*, 32, 2005, pp 187-202.
30. Levinson, D. Accessibility and Journey to Work, *Journal of Transport Geography*, 6(1), 1998, pp. 11-21.
31. Vandersmissen, M-H., Villeneuve, P. and Thériault, M. Analyzing Changes in Urban Form and Commuting Time, *The Professional Geographer*, 55(4), 2003, pp.446-463.
32. Jenkins 2005, Survival Analysis, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.176.7572&rep=rep1&type=pdf>
33. What is Duration Analysis? 2017. <http://jdemeritt.weebly.com/uploads/2/2/7/7/22771764/introduction.pdf>

34. Joly, I. Stability or Regularity of the Daily Travel Time in Lyon? – Application of a Duration Model, *International Journal of Transport Economics*, XXXIII, Vol. 3, 2006, pp. 369-400
35. Hensher, D. A. and F. L. Mannering. Hazard-Based Duration Models, and Their Application to Transport Analysis, *Transport Reviews*, Vol. 14, 1994, pp. 63-82.
36. Washington, S., Karlaftis, M. G. and F. L. Mannering. *Statistical and Econometric Methods for Transportation Data Analysis*. 2nd Edition, CRC Press, Taylor and Francis Group, 2011.
37. Bhat, C. R. A Hazard-based Duration Model of Shopping Activity with Nonparametric Baseline Specification and Non-parametric Control for Unobserved Heterogeneity. *Transportation Research B*, Vol. 30, 1996, pp. 189-207.
38. Komma, A. and S. Sirinivasan. Modelling Home-to-work Commute-Timing Decisions of Workers with Flexible Work Schedules. Transportation Research Board Annual Meeting, 2008.
39. Mohammadian, A. and S. T. Doherty. Modelling Activity Scheduling Time Horizon: Duration of Time between Planning and Execution of Pre-planned Activities, *Transportation Research A*, Vol. 40, 2006, pp. 475-490.
40. Zhong, M. and J. D. Hunt. Exploring Best-fit Hazard Function and Lifetime regression models for Urban Weekend Activities: Case Study. *Journal of Transportation Engineering*, ASCE, Vol. 136, No. 3, 2010, pp. 255-266
41. Oh, L. and J.W. Polak. Modelling Household Activity Duration Decisions: Accommodating Hierarchical Data structure, Selected proceedings of the 9th World Conference on Transport Research, Seoul, South Korea, 2002.
42. Schönfelder, S. and K. W. Axhausen. Analysing the rhythms of travel using survival analysis. Working paper submitted to TRB Washington DC - January 2001.
43. The population of the boroughs in the Metropolitan County of Tyne and Wear 2016. www.citypopulation.de/php/uk-tyneandwear.php
44. Local transport plan, *Tyne and Wear Household Travel Survey Report 2006/07*. <http://www.tyneandwearltp.co.uk/>. Accessed on Nov. 2008.
45. Ottewell, D., and Houlison, S., 2016, One of the UK's most congested roads outside of London is here in the North East, www.chroniclive.co.uk/news/north-east-news/one-uks-most-congested-roads-12038093
46. Citymetric report, 2016, Where are the UK's Most Congested Cities for Commuters? www.citymetric.com/transport/where-are-uks-most-congested-cities-commuters-2124
47. Kelly, M., 2017, Newcastle has some of the worst air pollution in the UK - so why do so few people care? www.chroniclive.co.uk/news/north-east-news/newcastle-worst-air-pollution-uk-12676181
48. Walker, J., Plan for North East congestion charge of £12.50 a day to cut air pollution, www.chroniclive.co.uk/news/north-east-news/plan-north-east-congestion-charge-11908922
49. DfT. *Regional Transport Statistics*, Department for Transport, UK, 2009
50. The great recession 2008-13, www.economicshelp.org/blog/7501/economics/the-great-recession
51. Northern way, 2009, City Relationships, Economic Linkages in Northern city regions, Tyne & Wear City Region, <http://www.theworkfoundation.com/assets/docs/tyne%20and%20wear.pdf>
52. Ettema, D. F., A. W. J. Borgers and H. J. P. Timmermans. Competing risk hazard model of activity choice, timing, sequencing and duration. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1493, Transportation Research Board of the National Academies, Washington, D.C., 1995, pp. 101-109.
53. Meyer, B. D., 1990, Unemployment Insurance and Unemployment Spells, *Econometrica*, Vol. 58, No. 4, 1990, pp. 757-782.
54. NCL travel plan 2008, www.ncl.ac.uk/estates/about/documents/NewcastleUniTPUpdate-Final.pdf
55. Delivering Cycling Improvements in Newcastle, A ten year strategy 2011-22, www.newcastle.gov.uk/sites/default/files/wwwfileroot/parking-roads-and-transport/cycling/af191248pd_cycling_brochure_2.pdf
56. Chen, W., Guo, A., W., and Blythe, P., 2017. Understanding Travellers' Satisfaction with and Attitudes Towards TfL Bus Service by Twitter, [//abstracts.aetransport.org/paper/index/id/4517/confid/20](https://abstracts.aetransport.org/paper/index/id/4517/confid/20)