

Segmenting travellers based on day-to-day variability in work-related travel behaviour

Travel needs for commute and business trips are complex and choices are not made based on the characteristics of individual trips, but instead based on the needs over weeks and months. For example, the cost per trip of commuting by bus varies depending upon the frequency of travel, and the cost of a monthly subway pass depends upon the number of zones visited during that period. Intrapersonal variability, namely the variation in an individual's travel behaviour from day to day, therefore shapes our transport choices and should influence service provision. Changes in working patterns such as increases in part time working, self-employment and tele-commuting challenge the traditionally held assumptions that work activities are fixed in time and space, thus making intrapersonal variability increasingly relevant. This research uses a data-driven approach to segment workers based on their work-related travel behaviour, including frequency of travel and both spatial and time of day intrapersonal variability. The analysis uses survey and seven day travel diary data for over 110,000 people collected over a 19 year period in England. Four groups of workers were identified: infrequent, spatially variable, temporally variable and regular travellers. These groups do not align closely with self-reported working arrangements such as self-employment or part time working. The group of regular travellers has decreased in size between 1998 and 2016 but remains the largest group, containing just under 60% of workers in 2016. Both the infrequent and spatially variable groups have grown over the same period. For a small but growing group of *workers*, a seven day diary is insufficient to understand their work-related transport needs as little or no work travel is recorded. These findings have implications for the design of public transport ticketing, the design of mobility as a service packages and the appraisal of congestion charging schemes.

Keywords: intrapersonal variability; commuting; working patterns; flexible working; multiday travel behaviour; business travel

1 Introduction

Over a quarter of the distance people travel in England is for commuting or is in the course of work (Department for Transport, 2018a). These trips occur disproportionately during the peak periods and contribute 56% of the trips between 7am and 8am on weekdays (Department for Transport, 2018b). It is crucial, therefore, to understand the needs underlying these trips when managing transportation networks and planning for the future. Whilst most research and practice characterises commuters and commuting trips based on average behaviour or a 'typical' day, this paper argues that measuring the frequency and day to day variability in an individual's behaviour is just as important when developing new policies and services.

The need to examine multiday travel behaviour has long been acknowledged (Pas and Koppelman, 1986; Huff and Hanson, 1986; Jones and Clarke, 1988). As well as just providing more information, collecting data over multiple days can also provide information about trips undertaken infrequently

and can provide important insights into intrapersonal variability, which is the variability in an individual's behaviour from day to day. Perhaps due to the increasing availability of multiday data, more researchers have focused on this area in recent years (Heinen and Chatterjee, 2015; Streit et al., 2015; McLeod et al., 2017; Kim et al., 2017; Heinen and Mattioli, 2017; Crawford et al., 2018). The majority of research into intrapersonal variability focuses on all trip types (or activity types) undertaken by an individual. In contrast, this research focuses on intrapersonal variability in commuting and business travel only.

Work is often assumed to be a fixed activity (Shen et al., 2013, p.2) which is "the base upon which much of a travel model system is built" (Miller, 2017, p.2). It is often assumed to be fixed in both time and space from day to day. This assumption has an impact on the design of services, for example people who need to visit clients during the working day may require integrated ticketing over a city or region. Estimated costs per trip by public transportation depend on both the cost of tickets and passes as well as the frequency of travel. Predicting traveller responses to planned or unexpected network disruptions also requires information about familiarity with the local network.

The assumption that individual commuters have highly predictable behaviour from day to day is increasingly questionable given that the nature of work is changing. In the UK, the employment rate for women is at the highest level since records began (Office for National Statistics, 2019a), the number of self-employed people increased by 15% between 1998 and 2016, and the service sector continues to grow (Office for National Statistics, 2019b). Flexible working has increased since the right to request flexible working was introduced in 2002 (Pyper, 2018), particularly part time and remote working. Technological developments have increased the "spatial fluidity" of work (Felstead, 2012, p.32), allowing some people to work at home or at other remote locations, including whilst travelling (Vilhelmson and Thulin, 2001). There has also been an increase in tele- and video-conferencing, although it is unclear the extent to which this has replaced travel (Vilhelmson and Thulin, 2001, p.1025). Regular or occasionally telecommuting is observed in many countries around the world (Eurofound and the International Labour Office, 2017, p.15).

As the nature of work has changed, so has the nature of commuting. Recent research in England identified a decrease in the total number of commuting trips, an increase in commuting distances and changes in the times of day at which people are commuting over the past two decades (Department for Transport, 2017). Whilst these changes aren't all necessarily due to changes in working arrangements, it is likely that such changes are a contributory factor. There is a substantial body of literature examining the impact of different types of working patterns on travel behaviour, including telecommuting (Mokhtarian et al., 2004; Choo et al., 2005; Haddad et al., 2009; Kim, 2017; Melo and de Abreu e Silva, 2017), self-employment (Gimenez-Nadal et al., 2018; Shin, 2019) and flexitime (He, 2013). Focusing on a specific working pattern can be problematic, however, as many terms do not have a universally agreed definition, such as 'teleworking' (Sullivan, 2003; Helminen and Ristimäki, 2007) and standard terms for working arrangements don't necessarily identify a homogeneous group of people. People working in the gig economy (Friedman, 2014) and remote workers (Sullivan, 2003) are two such 'groups' which contain subsets with very different characteristics in terms of job type and salary. Perhaps more importantly for this research, people within the same category may have very different work-related travel behaviour. As an example, Figure 1-1 shows three hypothetical working patterns for a part time employee working 21 hours per week. The number of days worked and the predictability in the timing of the travel varies across the three scenarios. This is not just the case for part-time workers, but also applies to people who occasionally work from home as this could be for full or part days (Haddad et al., 2009). In terms of transport planning for commuting and business trips, it would be more useful to identify people with

similar travel behaviour, irrespective of their working pattern, for example travel needs are the same for someone who does not work on Fridays as for someone who works at home all day on a Friday.

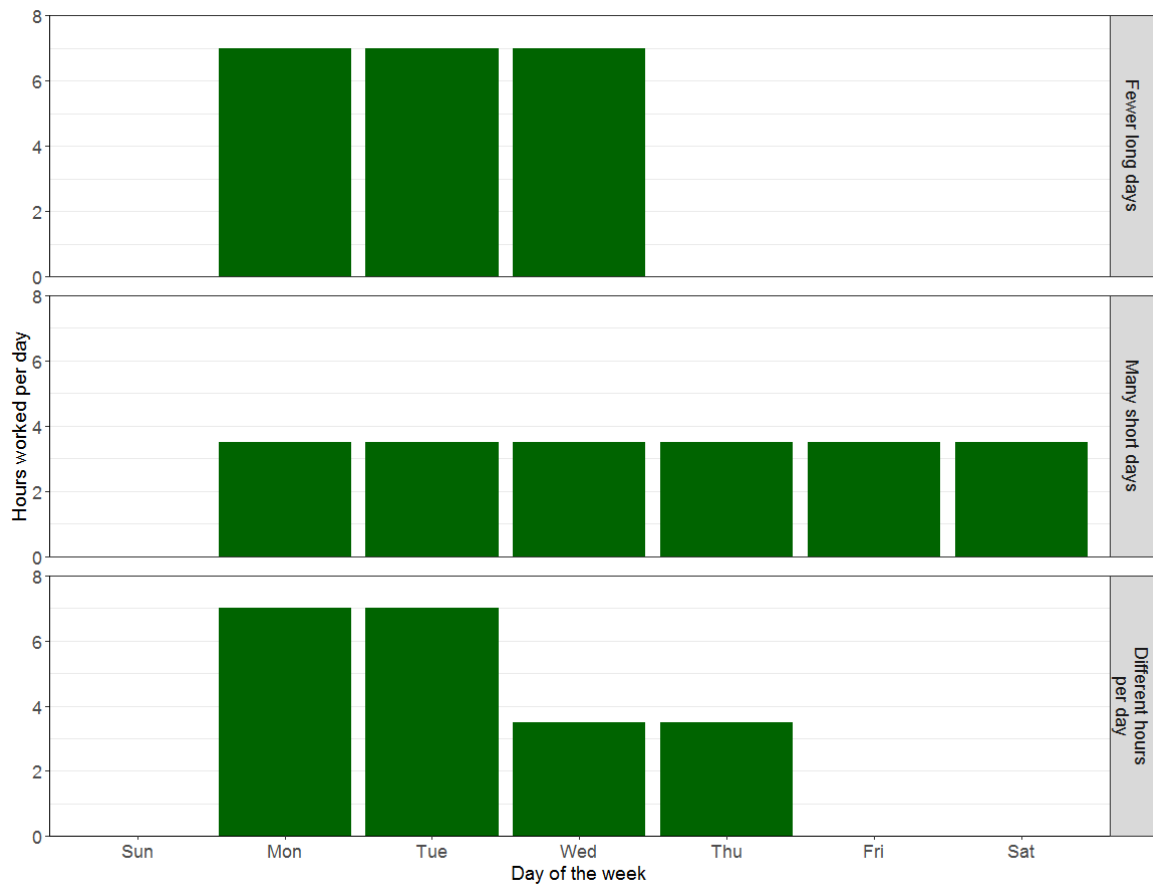


Figure 1-1: Hypothetical examples of working patterns consisting of 21 hours per week

Segmenting travellers based on multiple aspects of their travel behaviour can identify different groups of travellers with different needs which can inform policy and practice (Kieu et al., 2015; Goulet-Langlois et al., 2016; Crawford et al., 2018). Although other researchers have looked at segmenting travellers based on all trips, only Shen et al. (2013), to the author’s knowledge, have segmented travellers based on their commuting behaviour only. Shen et al. (2013) used GPS and travel diary data from 96 participants in Beijing for a seven day period to examine commuting intrapersonal variability based on time, space (measured in terms of the OD pair and also the route taken) and mode. Each of these characteristics were converted to a binary variable using pre-defined classifications of fixed and flexible behaviour. Each commuter was then assigned to one of seven commuter types based on whether they had fixed or flexible behaviour for each of the characteristic. They identified a substantial amount of variability, with only six respondents recording commuting trips which were ‘fixed’ in terms of all four repeated trip characteristics. They report that their findings “call into question the common presupposition that the commute trip is stable and fixed” (Shen et al., 2013, p.1).

This research builds on the work of Shen et al. (2013) by segmenting a much larger group of workers (over 118,000 people rather than 96) based on the frequency and intrapersonal variability in

commuting and business trips. The data was collected consistently over a 19 year period and therefore changes over time can be examined. The much larger dataset is as a result of using data from a nationwide survey, thus providing behavioural insights for transport providers across England. The scope of the current research is broader than previous research as it considers all trips to, from or during work as opposed to just commuting trips (between home and work only). This broader definition enables the segmentation process to consider all work-related trips which could impact upon mode choice and which could be affected by changes in working patterns and job types. Another difference in scope is that the current research focuses on the work-related aspects of variability from day to day so that the impacts of changes in working arrangements on travel can be observed and can be predicted given different policy options or expected trends. Variability in the places and times at which work is conducted (either by choice or necessity) are, therefore, considered but variability in route and mode choice are not.

The aim of the research is to identify different segments of the working population based on their multiday work-related behaviour to provide insights into travel needs. The approach is data-driven and no a priori assumptions about what constitutes 'fixed' or 'flexible' behaviour are made. The four research questions underpinning this paper are as follows:

- 1) Can different groups of workers be identified based on the frequency and intrapersonal variability in commuting and business trips?
- 2) To what extent can group membership be predicted based on data about work type from the National Travel Survey?
- 3) To what extent do other variables in the National Travel Survey, such as gender and age, influence the group a worker will belong to?
- 4) Are changes in the relative sizes of the groups observed over time? If so, what patterns emerge?

As described by Anable (2005, pp.66–67):

“Once groups are identified, it is possible to make predictions about their responses to various situations, marketing strategies and types of policy, to allow more creative and better-targeted policies to emerge.”

The segmentation of workers based on the variability in their work-related travel is therefore intended to provide a focus for policy makers and practitioners when considering the impact of policies on workers and when designing services (which may be transport related or not). The groups identified also provide an easy to explain framework which can be used in discussions with employers and with travellers. This new segmentation of travellers based on their day-to-day work-related travel behaviour seeks to provide inspiration for innovative solutions which do not necessarily just target traditional full time employees but which focus on current needs, perhaps through offering new transport services or incentivising different ways of working. The segmentation could also define user classes for day-to-day dynamical models or discrete choice modelling and could also provide inputs for activity-based models.

For the groups to be useful, it is important to be able to identify members without needing a seven day travel diary. This research, therefore, seeks to explore the characteristics of the groups of workers identified, in terms of working patterns and/or socio-demographic variables and the extent to which existing survey questions can identify the different work traveller types.

The paper is structured as follows. Section 2 describes the data used and the analytical techniques applied. In Section 3, the four groups of work travellers are described. The work and socio-

demographic characteristics of each group are described and logistic regression is used to estimate the impact of different characteristics on the likelihood of belonging to each cluster. Changes in the relative sizes of the groups over time are also shown. Section 4 discusses the results in terms of the research questions listed above. Section 5 concludes the paper.

2 Data and methodology

2.1 Data

An increasingly broad range of data sources can provide insights into travel behaviour. Multiday travel behaviour has been analysed using mobile phone data (Järv et al., 2014; Masso et al., 2019), Bluetooth data (Crawford et al., 2018), smartcard data (Kieu et al., 2015; Kim et al., 2017; Goulet-Langlois et al., 2018) and social media data (Zhang et al., 2017). Passively collected data has a disadvantage for the current research, however, as it does not provide the context required to separate out work travel (Pajević and Shearmur, 2017). Some researchers have proposed estimating trip purpose using a rule-based approach (Gong et al., 2014; Zou et al., 2018) or land use or point of interest data (Ermagun et al., 2017), but this risks further entrenching stereotypes about 'typical' working patterns and traditional workplaces.

This research instead uses travel diary data, where the respondent records the activity undertaken at each location. Due to the large amount of data collected consistently over a 19 year period, data from the National Travel Survey (NTS) in England was used. This large cross-sectional survey undertaken annually in England includes household and individual level surveys. Each household member also completes a seven day travel diary, including purpose, departure and arrival time, origin, destination, mode and distance travelled for each trip. The sample selection and data collection processes are described in detail in NatCen (2017).

Data from 1998 to 2016 was used (Department for Transport, 2015, no date) as prior to this employed people could not be easily identified in the data. Respondents were included in the analysis if they were at least 16 years old, employed, and if they completed the travel diary on two or more days. After restricting the sample based on these criteria, data from 118,194 people remained. The sample includes a slightly higher percentage of men than women (52% and 48% respectively) and 70% of the respondents were aged between 30 and 59. Overall, three quarters of respondents reported working full time and 13% reported being self-employed.

In collecting the data, the NTS uses the concept of a 'usual place of work'. This is the location where the respondent works on at least two consecutive days per week (NatCen, 2017, p.93). Respondents can only have one usual workplace and therefore if they work in multiple offices or have multiple jobs, the location used most frequently is recorded as their usual workplace. The percentage of people with a usual workplace in the sample decreased between 1998 and 2016, from 88% to 78%. The NTS also has 'in course of work' locations, which include secondary workplaces and locations visited for the purpose of meeting a client or providing a service to a customer. Commuting is defined in the NTS as "trips to a usual place of work from home, or from work to home" (Department for Transport, no date, p.11). Commuting trips therefore do not include travel to or from work involving non-trivial stops, for example to drop off a child at school or to buy groceries (Department for Transport, 2017). The NTS also records 'business trips', as shown in Figure 2-1. These trips involve an 'in course of work' location.

2.2 Method of analysis

2.2.1 Clustering workers

Despite the multi-dimensional nature of travel behaviour (Shen et al., 2013, p.2), the majority of the literature focuses on a single aspect of intrapersonal variability. Some research aims to *measure* intrapersonal variability, for example in mode choice (Crawford, 2019) or route choice (Arifin and Axhausen, 2012). Doherty (2006) measured both the temporal and spatial flexibility in activities undertaken and used these measures, amongst others, to classify different activities. Other research seeks to identify the factors which have a statistically significant impact on the type of variability under consideration, such as mode choice (Chatterjee et al., 2016), route choice (Li et al., 2005; Vacca and Meloni, 2015), or commute frequency (Helminen and Ristimäki, 2007). Research examining time of day variability for commuting trips has focused on identifying different components of the variation (Kitamura et al., 2006; Chikaraishi et al., 2009).

Cluster analysis has proved to be a suitable approach for examining multiple aspects of intrapersonal variability in previous research (Kieu et al., 2015; Crawford et al., 2018). Selecting the measures to use for clustering is key to identifying a meaningful segmentation. The three aspects of day-to-day variability explored in this research are trip frequency, spatial variability and temporal variability. As discussed in Section 1, these cover the day to day variability in where and when a person works which could be influenced by working patterns or arrangements. As it is the intrapersonal variability in the work behaviour which is of interest, day to day variability in transport mode and route are not considered as they are unlikely to be influenced by working arrangements.

Two types of measures are used to represent trip frequency: the number of days on which an individual visits a place of work and the number of work-related trips made on each of those days¹. The distinction is important as the former shows the spread of work-related travel through the week, whereas the latter provides a measure of the intensity of travel on those days.

Spatial analysis of this NTS data is limited by a lack of detailed origin and destination information due to privacy concerns. This research therefore considers spatial variability by separating trips to a usual workplace from trips to other work locations (which are not so fixed in space). A distinction is also made between commuting trips, which are between a fixed home location and a fixed usual workplace, and business trips, which are more varied. Figure 2-1 provides a summary of how work-related trip purposes are assigned based on the activities undertaken at the origin and the destination. The two types of trip frequency measures, namely days at work and number of trips per day, are therefore considered separately depending upon whether they relate to fixed locations (represented by the 'usual workplace' and commuting trips) or more variable ones (represented by 'in course of work' locations and business trips).

¹ Trips identified as 'a series of calls' were excluded as these do not involve spending a significant amount of time at any of the interim locations and thus they are more similar to a delivery schedule than travel for work. No weighting was required to account for 'drop off' during the diary week as the measure was trips per day.

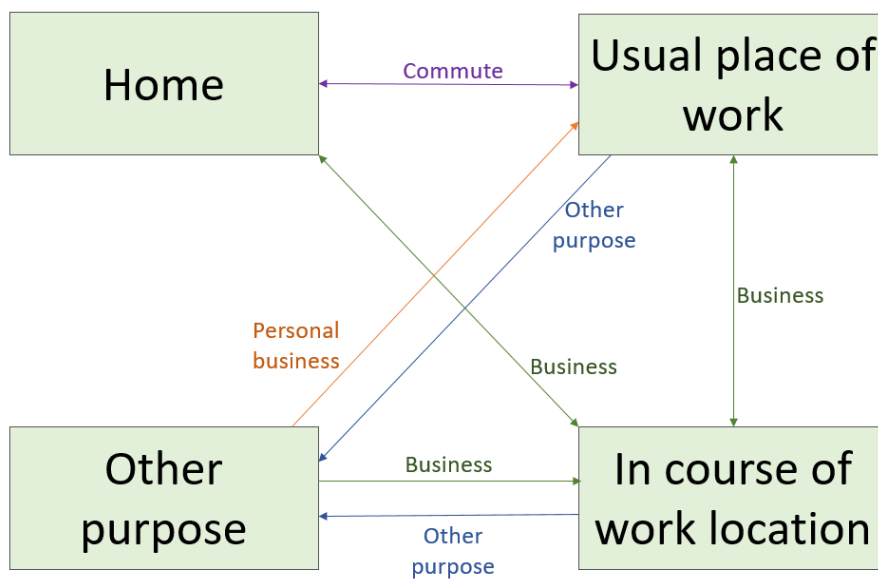


Figure 2-1: NTS trip purposes based on origin and destination activity purposes

Temporal, or time of day, variability is inextricably linked to the choice of origin and destination, as discussed in Shen et al. (2013, p.3). Using the NTS data, the only work-related trips which have consistent origins and destinations are commuting trips. In order to have a consistent basis for comparison, therefore, the departure times for commuting trips were used to calculate time of day variability. As commuting time variability can differ between the outbound and the inbound legs, both were considered as measures in the cluster analysis.

The six clustering variables used to identify workers with similar multiday work-related travel behaviour are listed in Table 2-1. These six measures were calculated for each respondent. The measures were then normalised before k-means clustering was applied using the statistical software R (R Core Team, 2019). The Elbow Method, which considers the number of clusters relative to the proportion of variability explained, was used to determine the number of clusters in the data.

Table 2-1: Clustering variables

Percentage of diary days the usual workplace is visited
Percentage of diary days an in course of work location is visited, but the usual workplace is not
Average number of commuting trips per day in their usual workplace
Average number of business trips per day in any workplace
Standard deviation of departure times from home to work (in minutes)
Standard deviation of departure times from work to home (in minutes)

Cluster analysis is sensitive to the data and the clustering variables used as inputs as well as the clustering algorithm used. To consider the impact of the choices made, sensitivity testing was undertaken, including repeating the cluster analysis using only data from people who completed all

seven days in the travel diary. Sensitivity testing using clustering methods other than k-means was also explored. Due to the size of the dataset, neither hierarchical clustering nor density based clustering methods were feasible. As an alternative, a two stage process including k-means clustering then hierarchical clustering was undertaken. This involved using k-means with a large number of clusters (in this case 1,000) and then using the centres of these clusters as an input for hierarchical clustering. By doing so, a small enough sample of observations is selected for hierarchical clustering to be feasible, whilst ensuring all of the data is represented. Results of the sensitivity analyses are discussed in Section 3.1.

2.2.2 Analysis of cluster characteristics

After clustering the workers based on the frequency and variability of their travel to, from and during work, the attributes of the different clusters were examined using profiling attributes (variables which were not involved in the clustering process) which link back to the research questions. The profiling attributes fell into two categories: those relating to working patterns and those relating to socio-demographic attributes.

In theory, multinomial regression could be used to examine the relationship between cluster membership and the various profiling attributes considered. A multinomial logit model can not be used, however, as the outcome variable describes cluster membership and therefore independence of irrelevant alternatives cannot be assumed. Multinomial probit regression could be used, in theory, but due to the large size of the dataset it was not possible to perform this analysis using a desktop computer². Instead, a separate logistic regression model was developed for each cluster, using the glm function in R (R Core Team, 2019), by estimating the parameters in the following equation:

$$\log_e \left(\frac{p_j}{1 - p_j} \right) = \beta_{0,j} + \sum_{i=1}^m \beta_{i,j} x_i$$

Where p_j is the probability that a randomly selected individual is in cluster j , $\beta_{0,j}$ is the intercept for the model for cluster j , $\beta_{i,j}$ is the coefficient for attribute i in the model considering membership of cluster j , and x_i is the i th attribute. Each model considers only membership of cluster j . Different specifications of the model (including different profiling attributes) were considered and then compared using the Akaike information criterion (AIC), with a lower AIC representing a better model. The same profiling attributes were included when modelling membership of each cluster so that the results would be comparable. This was somewhat limiting as only attributes which were available or which could be calculated across all clusters could be included in the models.

In order to apply logistic regression, all observations must be independent of one another. This assumption does not hold for the full dataset as each person within the sampled households completed a survey and travel diary and the working pattern of one person may be influenced by the working pattern of another person living in the same household, for example to satisfy childcare requirements. As in Heinen and Chatterjee (2015), this research randomly sampled one person from

² Initially the ‘mlogit’ package (Croissant, 2020) was used, but this was unable to obtain a solution on a desktop computer even for a sample of 10% of the data. The package ‘MNP’ (Imai and van Dyk, 2017) was also tried and this package was able to identify a solution on the full dataset, but there was insufficient memory on a desktop computer to run the procedure for determining the convergence diagnostics and therefore the results could not be used.

each household to reduce the risk of having individuals with dependent working patterns in the data. Another assumption of logistic regression is that there is little or no multicollinearity between the attributes included in the model. This was tested by calculating Variance Inflation Factor (VIF) values. None of the values were above 1.4 and therefore multicollinearity was not a concern for the models.

In this research, logistic regression has been used as a classification algorithm for the cluster membership and therefore, following the machine learning literature, performance measures arising from the confusion matrix have been used. After identifying the final logistic regression models, the data was randomly split into two components: a training set, containing 80% of the observations, and a test set, containing the remaining 20% of observations. Logistic regression models were estimated using the training data and then applied to the test data to predict whether or not an observation would belong to each cluster, using a 0.5 probability as the threshold for membership. The following performance metrics were then calculated for the model for each cluster, j , using the test data:

- accuracy (the proportion of observations which were correctly classified as either belonging to or not belonging to cluster j),
- recall (the proportion of observations in cluster j which were correctly classified as belonging to cluster j),
- precision (the proportion of observations predicted to be in cluster j which did actually belong to cluster j), and
- the area under the receiver operating characteristic (ROC) curve, known as the AUC, which provides an overall measure of model performance.

As the performance metrics will vary depending on the choice of training and testing sets, the process of randomly partitioning the data and then calculating the metrics was repeated 250 times. The standard deviations for all of the metrics were low (accuracy < 0.4%; precision < 4%; recall < 1.5%; AUC < 0.01) so only the average values across all 250 runs are reported.

3 Results

3.1 Overall

An analysis of data from 1998 to 2016 identified four clusters of workers. Table 3-1 contains the mean value for each of the six cluster variables for each group. The infrequent and the spatially variable work travellers, on average, travel to their usual workplace less than one day per week. These two clusters differ in terms of the travel undertaken to other locations for work. The spatially variable group visit other places of work on four days per week, on average, whereas the infrequent group have very little work-related travel to any location.

The temporally variable and the regular work travellers visit their usual workplace approximately five days per week on average. The temporally variable group show greater variability in the departure times of their commuting trips both to and from work. They also have 2.0 commuting trips, on average, per day at the usual workplace compared with 1.6 for regular commuters. This means that the temporally variable group are more likely to travel directly between home and work. The lower number of commuting trips for the regular group could be as a result of undertaking other activities directly before or after work, or staying overnight in a different location such as a second home as was observed in Helminen and Ristimäki (2007) and Shen et al. (2013).

Table 3-1: Average multiday work-related cluster characteristics (n=118,194)

Work travel group	Days per week at usual workplace	Additional days per week at any other workplace	Commuting trips per day at usual workplace	Business trips per day at any workplace	Variability in departure time (home to work) (mins)	Variability in departure time (work to home) (mins)	% of workers
Infrequent	0.6	0.1	0.1	0.1	0.1	0.1	20%
Spatially variable	0.6	3.9	0.3	2.1	3.4	3.6	8%
Temporally variable	5.0	0.1	2.0	0.1	225.6	226.5	12%
Regular	4.7	0.1	1.6	0.1	15.3	33.0	60%

By far the largest group is the ‘regular’ work travellers. This group does not necessarily consist of only ‘typical’ 9am to 5pm workers, however. Although they work consistent hours across the week, this analysis does not consider the length of the working day and so people who always work the same hours, whether that is a two hour or a twelve hour shift, are likely to appear in the same cluster. Also, the time of day analysis does not differentiate between day shifts and night shifts, provided that the departure times are consistent between days. The values in Table 3-1 are averages, so not all of the regular work travellers visited their usual workplace on five out of seven diary days. Of the people in this group who completed all seven days of the travel diary, 48% visited their usual place of work on exactly five days.

Sensitivity testing was undertaken by repeating the cluster analysis using only data from people who completed all seven days in the travel diary. This analysis resulted in four clusters, each with similar characteristics to the clusters in the original analysis. There were small differences in the percentage of workers assigned to each group, with fewer infrequent and more regular work travellers, but as there could be systematic differences between people who complete all seven days of the diary compared with people completing fewer days, the analysis including all people with at least two diary days was retained.

The other type of sensitivity analysis undertaken related to the type of clustering algorithm used. A two stage process including k-means clustering then hierarchical clustering resulted in three clusters which broadly consisted of the spatially variable workers, the temporally variable workers and then the regular and the infrequent commuters combined. The alternative method therefore produced some consistent findings but the original k-means clustering was deemed to be superior as it gave greater insights into travel behaviour by separating the regular from the infrequent work travellers.

3.2 Cluster characteristics

As described in Section 2.2.2, one person was randomly selected from each household and then logistic regression models were estimated for each of the four clusters. The list of independent variables included in the final model are shown in Table 3-3. Alternative models were developed to inform the final choice, including one containing the work-related independent variables only and

another where the binary variable representing living in London (or not) was replaced by a categorical variable representing all home regions. A consistent dataset was required for comparison and therefore people without region data were removed from the dataset, leaving a sample of 63,804 people for the regression analysis. As shown in Table 3-2, the models including age, sex and some kind of location variable have lower AIC values and therefore the models with only work variables were rejected. There is very little difference, in terms of the AIC, between the models including a binary indicator for people living in London and the models including all regions. The former was selected based on the principle of parsimony as it has fewer parameters.

Table 3-2: AIC for different model runs (n=63,804)

Cluster	Models with work-related variables only	Models with work-related variables plus age, sex and London indicator	Models with work-related variables plus age, sex and region
Infrequent	58,946	58,707	58,714
Spatially variable	29,479	29,464	29,462
Temporally variable	43,319	43,110	43,114
Regular	77,318	77,168	77,165

Table 3-3 includes the coefficients for the four final models (one for membership of each cluster). The figures in bold indicate coefficients which were statistically significant. Of the work-related attributes, five were significant for all four clusters: full/part time, home worker, work from home frequency, having a usual workplace, and working in professional or managerial occupations. Even after controlling for these, however, sex and age were statistically significant in all four models.

Table 3-3: Estimated coefficients within the logistic regression models (n=63,804)

	Infrequent	Spatially variable	Temporally variable	Regular
<i>Number of observations in cluster</i>	13,095	5,379	7,094	38,236
<i>Intercept</i>	-1.87***	-4.00***	-1.31***	0.71***
<i>Full time</i>	ref	ref	ref	ref
<i>Part time</i>	0.76***	-0.11*	-0.08*	-0.55***
<i>Employee</i>	ref	ref	ref	ref
<i>Self employed</i>	0.30***	0.33***	0.00	-0.38***
<i>Not a home worker</i>	ref	ref	ref	ref
<i>Home worker</i>	2.17***	1.92***	-1.85***	-2.83***
<i>Works from home less than once per week (or never)</i>	ref	ref	ref	ref
<i>Works from home at least once per week</i>	0.47***	0.65***	-0.45***	-0.65***
<i>Works in the same place on at least 2 consecutive days per week (has a 'usual' workplace)</i>	ref	ref	ref	ref
<i>Doesn't work in the same place on at least 2 consecutive days per week</i>	0.71***	2.35***	-1.02***	-1.27***
<i>Individual income less than £25,000</i>	ref	ref	ref	ref
<i>Individual income between £25,000 and £49,999</i>	0.01	0.11**	-0.03	-0.02
<i>Individual income of at least £50,000</i>	-0.02	0.29***	-0.38***	0.06
<i>Does not work in managerial or professional occupations</i>	ref	ref	ref	ref
<i>Works in managerial or professional occupations</i>	-0.20***	0.47***	-0.32***	0.13***
<i>Aged 20 or under</i>	ref	ref	ref	ref
<i>Age 21-59</i>	-0.14**	0.42**	-0.24***	0.17***
<i>Aged 60+</i>	-0.05	0.39**	-0.33***	0.15**
<i>Male</i>	ref	ref	ref	ref
<i>Female</i>	0.34***	-0.11**	-0.31***	-0.06**
<i>Lives outside London</i>	ref	ref	ref	ref
<i>Lives in London</i>	-0.19***	-0.06	-0.33***	0.30***

Values in bold are statistically significant

* represents $p < 0.05$, ** represents $p < 0.01$ and *** represents $p < 0.001$

ref represents the reference group

In Table 3-4, the coefficients have been transformed into odds ratios which show how much more or less likely a person is to belong to the specified cluster if they have the characteristic in the first column compared with people in the reference group (in column two). For example, part time workers are 2.14 times as likely to be in cluster 1 as full time workers, in other words the odds of part time workers being in cluster 1 is 114% higher than for full time workers.

Table 3-4: Odds ratios from each of the four cluster logistic regression models

	Reference group	Infrequent	Spatially variable	Temporally variable	Regular
Part time	Full time	2.14***	0.90*	0.92*	0.58***
Self employed	Employee	1.35***	1.40***	1.00	0.69***
Home worker	Not a home worker	8.72***	6.82***	0.16***	0.06***
Works from home at least once per week	Works from home less than once per week (or never)	1.59***	1.92***	0.64***	0.52***
Doesn't work in the same place on at least 2 consecutive days per week	Works in the same place on at least 2 consecutive days per week	2.04***	10.53***	0.36***	0.28***
Individual income between £25,000 and £49,999	Individual income less than £25,000	1.01	1.12**	0.97	0.98
Individual income of at least £50,000	Individual income less than £25,000	0.98	1.33***	0.68***	1.06
Works in managerial or professional occupations	Does not work in managerial or professional occupations	0.82***	1.60***	0.73***	1.13***
Age 21-59	Aged 20 or under	0.87**	1.53**	0.79***	1.19***
Age 60+	Aged 20 or under	0.95	1.48**	0.72***	1.16**
Female	Male	1.40***	0.90**	0.73***	0.94**
Lives in London	Lives outside London	0.83***	0.94	0.72***	1.34***

Values in bold are statistically significant

* represents $p < 0.05$, ** represents $p < 0.01$ and *** represents $p < 0.001$

All other things being equal, females are 40% more likely to be in the infrequent cluster than males. Table 3-5 shows the percentage of workers from the full sample used in Section 3.1 that belong to each of the four clusters. The table indicates that the increased likelihood resulted in the infrequent work travellers being the only group with a higher proportion of females than males. As might be expected from the low proportion of days in a workplace, part time workers are twice as likely to belong to the infrequent cluster as full time workers. Whilst this resulted in the infrequent cluster containing the highest percentage of part-time workers amongst the four clusters, the majority of people in the cluster (59%) report working full time. Similarly, self-employed workers were 35% more likely to belong to this cluster, but as there are far fewer self-employed people in the data, they only account for 23% of the workers in the infrequent cluster. Home workers³ are 772% more likely to belong to the infrequent cluster than people whose usual workplace is not at their home, although again due to a small number of homeworkers they are in the minority in this cluster as two

³ Home workers are defined as people who state that their home is their usual workplace.

thirds report having a usual workplace away from their home. Workers have higher odds of being in the infrequent or the temporally variable cluster if they are aged 20 or younger, all other things being equal.

Table 3-5: Characteristics of workers assigned to each of the four clusters

	Infrequent (n = 23,680)	Spatially variable (n = 9,580)	Temporally variable (n = 13,699)	Regular (n = 71,235)
Female	57%	35%	44%	47%
Part time	41%	21%	25%	21%
Self-employed	23%	34%	8%	7%
Home worker	14%	12%	1%	1%
Works from home at least once per week	9%	19%	3%	4%
In professional or managerial occupations	41%	58%	34%	44%
Earns under £25k per year	71%	53%	71%	65%
Lives in London	12%	14%	10%	15%
Lives in an urban area ⁴	77%	75%	81%	83%
Aged 20 or under	6%	2%	8%	5%
Aged 60 or over	12%	12%	8%	8%

The spatially variable cluster contained the smallest percentage of workers in the original sample. All other things being equal, workers with salaries above £25,000 are more likely to be in this cluster than those with lower incomes. People working in professional or managerial roles were 60% more likely to be in this cluster than people working in other types of jobs, all else being equal. Part time workers had a *lower* likelihood of belonging to this cluster than full time workers whereas self-employed workers were 40% *more* likely to be in the spatially variable cluster than employees. Females have a lower likelihood of being in this cluster than males, all else being equal, resulting in the group being dominated by males (65% of the workers assigned to the spatially variable cluster). Home workers are 582% more likely to belong to the spatially variable clusters than people whose usual workplace is not at their home. The other very large impact comes from *not* having a 'usual workplace' as the likelihood of being in the spatially variable cluster is over 10 times that of people who do have a 'usual workplace'. It was not possible to include the urban/rural indicator in the regression models as the data is only available for surveys undertaken since 2002, but Table 3-5 shows that this cluster has the lowest percentage of workers who live in urban areas, although 75% do live in urban areas.

The vast majority of people in the temporally variable and the regular work-travel groups are employees who have a usual workplace and rarely work from home. Workers with salaries of at least £50,000 are less likely to be in the temporally variable cluster than those with incomes below £25,000 and, similarly, people working in professional or managerial roles are less likely to belong to this cluster which is defined by frequent but temporally variable commuting. Females are less likely than males to belong to the temporally variable cluster, all other things being equal. Whilst the odds ratio for females versus males is lowest for the temporally variable cluster, the spatially variable

⁴ This classification is not available for data prior to 2002

cluster has the lowest percentage of females assigned to it in total, perhaps because other factors such as self-employment and home working are not evenly distributed between genders.

The regular cluster contains many workers with stereotypically ‘traditional’ working patterns. Membership of this cluster is *less* likely for part time workers (relative to full time workers), self-employed people (relative to employees), home workers (relative to those who do not use their home as their usual workplace), people without a usual workplace (relative to those who do) and people who work from home at least once per week (relative to those who do not). The odds ratio for part time relative to full time working is lower than for the temporally variable cluster, perhaps suggesting that some workers in the temporally variable cluster may work some partial days. Workers living in London have 34% higher likelihood of belonging to the regular cluster than those who live outside of London.

The performance metrics for the final model are shown in Table 3-6. The AUC and accuracy values suggest that the models are performing relatively well. To put this in context, however, if the models for the infrequent, spatially variable and temporally variable clusters assigned no workers to these clusters and the model for the regular cluster assigned all workers to this cluster, the accuracy for the models would be 80%, 92%, 88% and 60%, respectively. The model for the temporally variable group in fact appears to work in this way, as typically no workers from test sets were predicted to belong to the group. The values for recall, which represents the percentage of people in this cluster who were correctly predicted to be in the cluster, show that for the models for the infrequent, spatially variable and temporally variable groups the accuracy is based on correctly predicting that certain workers *are not* in the group, whereas for the regular group much of the accuracy comes from correctly predicting which workers *are* in this cluster. The performance metrics suggest that while the profiling attributes can be helpful in predicting cluster membership in some cases, there are factors influencing the work-related travel types which are not currently being collected either in the working pattern or the socio-demographic data in the NTS.

Table 3-6: Average performance metrics for the final model specification over 250 runs (80% training data; 20% test data)

	Average accuracy	Average precision	Average recall	Average AUC
Infrequent	81%	64%	15%	0.70
Spatially variable	91%	46%	7%	0.81
Temporally variable	89%	N/A	0%	0.63
Regular	69%	68%	90%	0.69

3.3 Changes over time

In the results presented above, data from 1998 to 2016 was combined. Figure 3-1 shows how the relative sizes of the four clusters have changed over that period. The infrequent and the spatially variable clusters have both grown, with their combined size rising from 21% in 1998 to 32% in 2016. The temporally variable cluster has decreased in size from 16% of working respondents in 1998 to 9% in 2016. The regular commuting cluster remains by far the largest cluster, although its relative size has decreased from 63% to 59%.

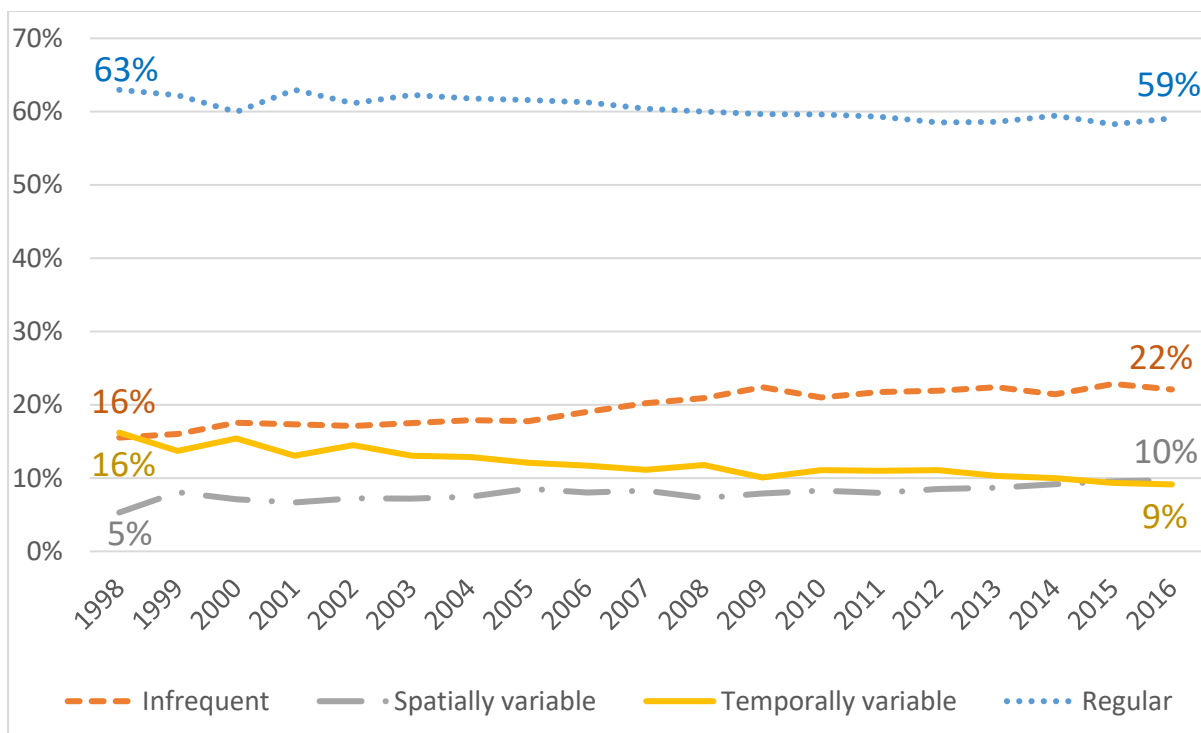


Figure 3-1: Work travel clusters over time

3.4 Travel behaviour

It could be hypothesised that there is a relationship between the working patterns of an individual and the distance between their home and their usual workplace. In the publicly available NTS data, information on distance to workplace is not given directly. The distance can be estimated from the diary data, but only for workers who record a commute trip (directly between home and the usual workplace) during the week. As expected, based on the low number of commute trips recorded, the commute distance is only available for 17% of infrequent travellers and 24% of spatially variable work travellers. Commute distance is available for over 99% of temporally variable and regular commuters, however. On average, regular commuters have slightly longer commutes than the temporally variable group (median one-way distance 6 miles and 4 miles, respectively).

The mode choice for work trips can be examined in two ways using the NTS data. Firstly, each person is asked within the survey how they usually travel to work. Secondly, the actual mode choice for commuting and other business trips made during the surveyed week are recorded in the travel diaries. When examining the usual commuting mode in the survey data, car, either as a driver or a passenger, is the most common mode used in all groups (Table 3-7). Car use was highest (80%) in the spatially variable cluster and 14% of the people in this group report being the main driver of a company car, compared with 3-4% for the other three groups. Public transport usage is slightly higher for the regular work travellers and the infrequent group report the highest levels of walking to work. The composition of public transport users is not the same across groups, as the spatially variable travellers were more likely to travel by train, underground, metro or tram, whereas the temporally variable commuters were more likely to be bus users.

Table 3-7: Reported usual means of travel to work (NTS individual survey)

Usual mode of travel to work	Infrequent	Spatially variable	Temporally variable	Regular
Car	64%	80%	73%	69%
Public transport	12%	11%	10%	18%
Walk	19%	5%	8%	8%
Cycle	2%	2%	5%	4%
Other	3%	2%	3%	2%

In the travel diary data, 75% of the temporally variable commuters use the same main mode for all commuting trips in the given week. The equivalent figure for the regular commuters is 84%.

4 Discussion

The four research questions presented in Section 1 will now be discussed in order.

4.1 Can different groups of workers be identified based on the frequency and intrapersonal variability in commuting and business trips?

Whilst Section 3.1 demonstrated that four clusters or groups of workers can be identified, these groups are only useful if they identify different transport needs. For the infrequent group, the seven day travel diaries provide insufficient information to inform the design of services to meet their needs. These people stated that they are employed but did not record any work-related travel in the given week and there are a number of reasons why this might have been the case. It could be because the person works from or very near to their home, was on holiday during the survey week or the person was not scheduled to work that week, for example due to a 'zero hour' contract, working arrangements such as term-time working or other types of leave such as maternity leave. Indeed, a higher percentage of people who completed the travel diary during school holiday periods reported no work-related travel compared with people who completed the diaries during term time (23% and 13% respectively). Another possible explanation is that participants only record personal travel in the NTS. Personal travel is defined as "travel for private purposes or for work or education, provided the main reason for the trip is for the traveller himself or herself to reach the destination" (Department for Transport, no date, p.9). Trips which are undertaken for the purpose of transporting goods or people are therefore excluded. This cluster may, therefore, contain a very heterogeneous group in terms of their work-related transport needs and therefore further research is required to explore their requirements.

The spatially variable work travellers do not have a usual workplace which provides an anchor point in their lives and therefore they need a flexible means of transport. This may partially explain the high proportion of work trips made by car. The high percentage of self-employed people in this group may also play a role, as this group may be more likely to take products or equipment with them when they travel or there may be a perceived need to demonstrate the status associated with arriving by car. Another potential explanation is that the car use is related to the higher individual incomes of people in this group. Whilst demand responsive transport services could provide a more flexible shared transport solution for some of these workers, such services are not typically designed to carry equipment or stock so the types of services offered may need to be reconsidered. Similarly,

some of these workers may require their vehicle to remain close by so that they can access equipment which is stored inside.

The temporally variable group tend to travel directly between their home and a fixed work location many times per week. This group would therefore be good candidates for public transport use, although only 10% usually travel to work by public transport. They have high time of day variability but it is unclear whether the departure times on each day are known in advance. This research has deliberately used the term 'variability' in terms of commuting behaviour in contrast to the term 'flexibility' used by some other researchers, for example Streit et al. (2015) and Shen et al. (2013) (following Doherty (2006)). Variability in the commute does not necessarily mean that the *commuter* has the power to choose where and when they work (Felstead, 2012). These temporally variable workers may be free to leave work whenever they please, for example to fit with a scheduled public transport service, or they may have to stay as long as work needs dictate (potentially until late at night). In terms of Hägerstrand's (1970) space-time geography, the variability observed in space and/or time could mean that these activities are fairly flexible (as was assumed in Doherty (2006)) or it could mean that the work activity is fixed but that the constraints differ from day to day. These two possible explanations suggest very different transport needs so further research is required to understand the degree of spatial and temporal flexibility of activities, for example using the method demonstrated in Shen et al. (2015). The frequency of services and the availability of reliable real time travel information may be influential in the travel choices of this group or on-demand services may be more suitable.

The regular work travellers have the highest percentage of public transport use at 18%, but the majority use a car to get to work. Also, the higher rate of public transport use may be influenced by the higher likelihood of people living in London belonging to this cluster as opposed to the perception that public transport better meets their needs. This group travels to the same location at approximately the same time each day, so the travel diaries provide valuable data about the services they require. Two further aspects need to be considered, however. Firstly, not all members of this group travel to work on five days per week. In total, 29% travelled to their usual workplace on three or four days during the diary week and therefore they would have had a higher cost per trip than people working five days per week if purchasing weekly or monthly travel passes. Also, fewer people in this group travel directly between home and work than in the temporally variable group. The types of locations visited before or after work and the requirements of such travel, including carrying shopping or escorting a family member, requires further investigation to determine whether it is the chained trips rather than the trips to and from work which are less attractive by public transport.

4.2 To what extent can group membership be predicted based on data about work type from the National Travel Survey?

In Section 3.2, we observed that many of the attributes relating to working patterns, such as part time working, were related to cluster membership. The models containing only the working pattern attributes, however, were rejected in favour of models which also included age, sex and location attributes, suggesting that the working pattern attributes were useful but not sufficient to explain cluster membership.

Previous research has highlighted the heterogeneity in workers classified as self-employed (Gimenez-Nadal et al., 2018), mobile workers (Pajević and Shearmur, 2017), occasional home workers (Haddad et al., 2009) and people working in the 'gig economy' (Friedman, 2014). In the

current research we have highlighted the heterogeneity in the weekly work *travel* behaviour of part time and self-employed workers particularly. Part time workers constitute a high percentage of the infrequent commuters (41%), but due to differences in the sizes of the clusters, only a third of the part time workers in the sample are allocated to this cluster. Half of the part time workers in the sample are, in fact, assigned to the regular commuter cluster. Similarly, there are much higher rates of self-employment within the infrequent and spatially variable clusters, although these two clusters account for only 57% of the self-employed people in the sample. Therefore, whilst certain working patterns can increase the odds of belonging to a specific cluster, knowing working patterns alone is not sufficient to predict frequency and intrapersonal variability in work-related travel behaviour. This raises the question of how useful standard questions, about full or part time working for example, are when collecting data about travel and transport needs.

4.3 To what extent do other variables in the National Travel Survey, such as gender and age, influence the group a worker will belong to?

There were statistically significant coefficients relating to age group in the models for all four clusters. Even after accounting for part time working and all of the other attributes, workers aged 20 or under were more likely to be in the infrequent and the temporally variable clusters.

Females are 40% more likely to be in the infrequent cluster than males, all other things being equal. This highlights the need to understand the workers in this cluster. If it is the case that many of these women were not working during that particular week but they do undertake work-related travel in other weeks, then their needs are being under-represented in the data currently collected and the services designed based on that data.

Whilst age, gender and living in London were selected for inclusion in the final model, it is clear from Table 3-6 that the data currently collected in the survey part of the NTS is insufficient to be able to determine the work-related travel needs of individuals. The additional information required could relate to working patterns or job type, but could also be socio-demographic or geographic data that is not currently available.

4.4 Are changes in the relative sizes of the groups observed over time? If so, what patterns emerge?

Figure 3-1 demonstrated that the relative cluster sizes have been changing over time. The two work traveller clusters which are growing in size have a higher proportion of workers aged 60 or over than the groups which are shrinking. Given that the working population in England is ageing (Ross, 2010), these trends may well continue. This issue is not limited to England as many countries have increased or are increasing the state pension age or have reduced the value of state pensions (Cebulla et al., 2007, p.850). The growth in the infrequent cluster may also be due to an increase in women entering the labour market as they are more likely to belong to this cluster than men, all else being equal. Of the two groups which have been increasing, we have very little information about the work-related travel behaviour for one and the other is heavily car dominated. The two groups which are shrinking over time are the temporally variable and the regular clusters. These are the workers with the most predictable behaviour and for whom many traditional public transport services were designed. This strongly implies that collecting the right data and developing suitable policies now are crucial if we are to meet future work-related transport needs.

5 Conclusions

In this paper, data from seven day travel diaries in England collected between 1998 and 2016 were used to identify different types of workers based on the predictability of their travel behaviour over multiple days. The results indicate that there are four distinct groups with different spatial and/or temporal working patterns, and thus different transport needs. The two groups travelling most often to a 'usual workplace' have been decreasing over time.

The National Travel Survey provides a wealth of information about work travel but, as with any source of data, it has limitations. The NTS data which is publicly available does not include detailed geographical data and as discussed in Section 2.2, it was not possible to break down work activities undertaken in 'in course of work' locations. These could involve repeated trips to the same location or to vastly different geographic areas. These different scenarios impact upon the transport options available. Further work could therefore involve the augmentation of travel diary data with GPS tracker data (as in Shen et al. (2013)), although there would be cost and privacy implications.

This research highlights four distinct groups of workers in England with different transport needs relating to their work. All of the groups are sufficiently large to warrant consideration when designing services or policies which will have, or aim to have, an impact on commuting or business travel. Any intervention designed solely based on the 'traditional' regular worker is targeted at 59% of workers in 2016 at most. The other three groups combined are more likely to contain females, part time workers, self employed people and those who either always or occasionally work from home. It is essential that these other types of workers are explicitly considered when transportation or work policies are being developed so that they are not disadvantaged.

Questions are raised about the suitability of data collected about work and travel. The increasing proportion of people recording very little or no work-related travel in the diary indicates that something is missing. This could be an issue with the NTS question asking people whether they work, or it could be a limitation of collecting just one week of data. Either way, additional research is required to understand this group of people who are currently not represented in the work travel data. The NTS also does not collect information about work undertaken whilst at home or whilst travelling and therefore cannot be used to examine how work is spread across the week. The usefulness of standard terms such as 'full time', 'part time' and 'self-employed' in travel surveys is also brought into question as the groups are highly heterogeneous in terms of travel behaviour.

Previous work has highlighted that the nature of work and the nature of commuting at an aggregate level are changing. This research has identified that the predictability of travel to, from and in the course of work is also changing. As the multiday transport needs of workers change, so must the transport networks and services. Changes in traveller needs influence all aspects of transportation, including fare structures and real time information as well as the development of new services and infrastructure.

6 Acknowledgements

See title page

7 References

- Anable, J. (2005) 'Complacent Car Addicts' or 'Aspiring Environmentalists'? Identifying travel behaviour segments using attitude theory. *Transport Policy*. 12 (1), pp. 65–78.
- Arifin, Z.N. and Axhausen, K.W. (2012) Investigating Commute Mode and Route Choice Variability in Jakarta using multi-day GPS Data. *International Journal of Technology*. 1 pp. 45–55.
- Cebulla, A., Butt, S. and Lyon, N. (2007) Working beyond the state pension age in the United Kingdom: the role of working time flexibility and the effects on the home. *Ageing and Society*. 27 (06), pp. 849–867. [Accessed 27 February 2019].
- Chatterjee, K., Clark, B. and Bartle, C. (2016) Commute mode choice dynamics: Accounting for day-to-day variability in longer term change. *European Journal of Transport and Infrastructure Research*. 16 (4), pp. 713–734.
- Chikaraishi, M., Fujiwara, A., Zhang, J. and Axhausen, K.W. (2009) Exploring Variation Properties of Departure Time Choice Behavior by Using Multilevel Analysis Approach. *Transportation Research Record*. 2134 (1), pp. 10–20.
- Choo, S., Mokhtarian, P.L. and Salomon, I. (2005) Does telecommuting reduce vehicle-miles travelled? An aggregate analysis for the U.S. *Transportation*. 32 pp. 37–64.
- Crawford, F. (2019) Variability in commuting mode in England *Analyses from the National Travel Survey* [online].
- Crawford, F., Watling, D.P. and Connors, R.D. (2018) Identifying road user classes based on repeated trip behaviour using Bluetooth data. *Transportation Research Part A: Policy and Practice*. 113 pp. 55–74.
- Croissant, Y. (2020) *mlogit: Multinomial Logit Models*. URL <https://cran.r-project.org/package=mlogit> [Accessed 29 May 2020].
- Department for Transport (2017) *Commuting trends in England: 1988-2015* [online] (November).
- Department for Transport (2015) *National Travel Survey, 1995-2001, 3rd Edition*. UK Data Service. SN: 6108.
- Department for Transport (no date) *National Travel Survey, 2002-2016, 12th Edition*. UK Data Service. SN: 5340.
- Department for Transport (no date) *National Travel Survey 2016: Notes and Definitions*. URL https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/632914/nts-2016-notes.pdf [Accessed 24 May 2019b].
- Department for Transport (2018a) *NTS0403: Average number of trips, miles and time spent travelling by trip purpose: England*. URL <https://www.gov.uk/government/statistical-data-sets/nts04-purpose-of-trips> [Accessed 31 May 2019].
- Department for Transport (2018b) *Table NTS0502: Trip start time by trip purpose (Monday to Friday only): England, 2013/17, National Travel Survey 2017* [online].
- Doherty, S.T. (2006) Should we abandon activity type analysis? Redefining activities by their salient attributes. *Transportation*. 33 (6), pp. 517–536.
- Ermagun, A., Fan, Y., Wolfson, J., Adomavicius, G. and Das, K. (2017) Real-time trip purpose prediction using online location-based search and discovery services. *Transportation Research Part C: Emerging Technologies*. 77 pp. 96–112.

- Eurofound and the International Labour Office (2017) *Working anytime, anywhere: The effects on the world of work* [online].
- Felstead, A. (2012) Rapid change or slow evolution? Changing places of work and their consequences in the UK. *Journal of Transport Geography*. 21 pp. 31–38.
- Friedman, G. (2014) Workers without employers: shadow corporations and the rise of the gig economy. *Review of Keynesian Economics*. 2 (2), pp. 171–188.
- Gimenez-Nadal, J.I., Molina, J.A. and Velilla, J. (2018) The commuting behavior of workers in the United States: Differences between the employed and the self-employed. *Journal of Transport Geography*. 66 pp. 19–29.
- Gong, L., Morikawa, T., Yamamoto, T. and Sato, H. (2014) Deriving Personal Trip Data from GPS Data: A Literature Review on the Existing Methodologies. *Procedia - Social and Behavioral Sciences*. 138 pp. 557–565.
- Goulet-Langlois, G., Koutsopoulos, H.N. and Zhao, J. (2016) Inferring patterns in the multi-week activity sequences of public transport users. *Transportation Research Part C: Emerging Technologies*. 64 pp. 1–16.
- Goulet-Langlois, G., Koutsopoulos, H.N., Zhao, Z. and Zhao, J. (2018) Measuring Regularity of Individual Travel Patterns. *IEEE Transactions on Intelligent Transportation Systems*. 19 (5), pp. 1583–1592.
- Haddad, H., Lyons, G. and Chatterjee, K. (2009) An examination of determinants influencing the desire for and frequency of part-day and whole-day homeworking. *Journal of Transport Geography*. 17 (2), pp. 124–133.
- Hägerstrand, T. (1970) What about people in regional science? In: *Papers of the Regional Science Association, Vol. 24*. 1970 pp. 7–21.
- He, S.Y. (2013) Does flexitime affect choice of departure time for morning home-based commuting trips? Evidence from two regions in California. *Transport Policy*. 25 pp. 210–221.
- Heinen, E. and Chatterjee, K. (2015) The same mode again? An exploration of mode choice variability in Great Britain using the National Travel Survey. *Transportation Research Part A: Policy and Practice*. 78 pp. 266–282.
- Heinen, E. and Mattioli, G. (2017) Does a high level of multimodality mean less car use? An exploration of multimodality trends in England. *Transportation*.
- Helminen, V. and Ristimäki, M. (2007) Relationships between commuting distance, frequency and telework in Finland. *Journal of Transport Geography*. 15 (5), pp. 331–342.
- Huff, J.O. and Hanson, S. (1986) Repetition and Variability in Urban Travel. *Geographical Analysis*. 18 (2), pp. 97–114.
- Imai, K. and van Dyk, D. (2017) *MNP: R Package for Fitting the Multinomial Probit Model*. R package version 3.1-0. URL <https://cran.r-project.org/package=MNP> [Accessed 29 May 2020].
- Järv, O., Ahas, R. and Witlox, F. (2014) Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records. *Transportation Research Part C: Emerging Technologies*. 38 pp. 122–135.
- Jones, P. and Clarke, M. (1988) The significance and measurement of variability in travel behaviour. *Transportation*. 15 (1), pp. 65–87.

- Kieu, L.M., Bhaskar, A. and Chung, E. (2015) Passenger segmentation using smart card data. *IEEE Transactions on Intelligent Transportation Systems*. 16 (3), pp. 1537–1548.
- Kim, J., Corcoran, J. and Papamanolis, M. (2017) Route choice stickiness of public transport passengers: Measuring habitual bus ridership behaviour using smart card data. *Transportation Research Part C: Emerging Technologies*. 83 pp. 146–164.
- Kim, S.-N. (2017) Is telecommuting sustainable? An alternative approach to estimating the impact of home-based telecommuting on household travel. *International Journal of Sustainable Transportation*. 11 (2), pp. 72–85.
- Kitamura, R., Yamamoto, T., Susilo, Y.O. and Axhausen, K.W. (2006) How routine is a routine? An analysis of the day-to-day variability in prism vertex location. *Transportation Research Part A: Policy and Practice*. 40 (3), pp. 259–279.
- Li, H., Guensler, R. and Ogle, J. (2005) Analysis of Morning Commute Route Choice Patterns Using Global Positioning System–Based Vehicle Activity Data. *Transportation Research Record*. 1926 (1), pp. 162–170.
- Masso, A., Silm, S. and Ahas, R. (2019) Generational differences in spatial mobility: A study with mobile phone data. *Population, Space and Place*. 25 (2), pp. e2210.
- McLeod, F.N., Cherrett, T.J., Box, S., Waterson, B.J. and Pritchard, J.A. (2017) Using automatic number plate recognition data to investigate the regularity of vehicle arrivals. *European Journal of Transport and Infrastructure Research*. 17 (1), pp. 86–102.
- Melo, P.C. and de Abreu e Silva, J. (2017) Home telework and household commuting patterns in Great Britain. *Transportation Research Part A: Policy and Practice*. 103 pp. 1–24.
- Miller, E.J. (2017) Modeling the Demand for New Transportation Services and Technologies. *Transportation Research Record: Journal of the Transportation Research Board*. 2658 (2658), pp. 1–7.
- Mokhtarian, P.L., Collantes, G.O. and Gertz, C. (2004) Telecommuting, residential location, and commute-distance traveled: Evidence from State of California employees. *Environment and Planning A*. 36 (10), pp. 1877–1897.
- NatCen (2017) *National Travel Survey 2016: Technical Report*. URL https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/632910/nts-technical-report-2016.pdf [Accessed 24 May 2019].
- Office for National Statistics (2019a) *Female employment rate (aged 16 to 64, seasonally adjusted)*. URL <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/timeseries/lf25/lms> [Accessed 30 May 2019].
- Office for National Statistics (2019b) *Long-term trends in UK employment: 1861 to 2018*. URL <https://www.ons.gov.uk/economy/nationalaccounts/uksectoraccounts/compendium/economicreview/april2019/longtermtrendsinukemployment1861to2018> [Accessed 30 May 2019].
- Pajević, F. and Shearmur, R.G. (2017) Catch Me if You Can: Workplace Mobility and Big Data. *Journal of Urban Technology*. 24 (3), pp. 99–115.
- Pas, E.I. and Koppelman, F.S. (1986) An examination of the determinants of day-to-day variability in individuals' urban travel behavior. *Transportation*. 13 (2), pp. 183–200.
- Pyper, D. (2018) *Flexible working*. URL <https://researchbriefings.files.parliament.uk/documents/SN01086/SN01086.pdf> [Accessed 29

May 2020].

- R Core Team (2019) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.r-project.org/>
- Ross, D. (2010) Ageing and work: An overview. *Occupational Medicine*. 60 (3), pp. 169–171.
- Shen, Y., Chai, Y. and Kwan, M.-P. (2015) Space–time fixity and flexibility of daily activities and the built environment: A case study of different types of communities in Beijing suburbs. *Journal of Transport Geography*. 47 pp. 90–99.
- Shen, Y., Kwan, M.-P. and Chai, Y. (2013) Investigating commuting flexibility with GPS data and 3D geovisualization: a case study of Beijing, China. *Journal of Transport Geography*. 32 pp. 1–11.
- Shin, E.J. (2019) Self-employment and travel behavior: A case study of workers in central Puget Sound. *Transport Policy*. 73 pp. 101–112.
- Streit, T., Allier, C.-E., Weiss, C., Chlond, B. and Vortisch, P. (2015) Changes in Variability and Flexibility of Individual Travel in Germany: Trends and Drivers. *Transportation Research Record: Journal of the Transportation Research Board*. 2496 pp. 10–19.
- Sullivan, C. (2003) What’s in a name? Definitions and conceptualisations of teleworking and homeworking. *New Technology, Work and Employment*. 18 (3), pp. 158–165.
- Vacca, A. and Meloni, I. (2015) Understanding Route Switch Behavior: An Analysis Using GPS Based Data. *Transportation Research Procedia*. 5 pp. 56–65.
- Vilhelmson, B. and Thulin, E. (2001) Is Regular Work at Fixed Places Fading Away? The Development of ICT-Based and Travel-Based Modes of Work in Sweden. *Environment and Planning A: Economy and Space*. 33 (6), pp. 1015–1029.
- Zhang, Z., He, Q. and Zhu, S. (2017) Potentials of using social media to infer the longitudinal travel behavior: A sequential model-based clustering method. *Transportation Research Part C: Emerging Technologies*. 85 pp. 396–414.
- Zou, Q., Yao, X., Zhao, P., Wei, H. and Ren, H. (2018) Detecting home location and trip purposes for cardholders by mining smart card transaction data in Beijing subway. *Transportation*. 45 (3), pp. 919–944.