DTU Library

Considerations on a survey method to collect long distance travel data when only asking about the two latest journeys

Christensen, Linda

Published in:
Transportation Research Procedia

Link to article, DOI:
10.1016/j.trpro.2018.10.024

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 32 (2018) 553–562

Transportation Research Procedia

www.elsevier.com/locate/procedia

International Steering Committee for Transport Survey Conferences

Considerations on a survey method to collect long distance travel data when only asking about the two latest journeys

Linda Christensen*

Technical University of Denmark, DTU Management, Bygningstorvet 116V, 2800 Lyngby, Denmark

Abstract

The paper describes a new Danish survey about journeys abroad with overnight stay(s). The survey is used to investigate if it is possible to collect long distance travel data by only asking for information about the latest journey. It is shown that it is not enough to include the latest journey, and that it is necessary to ask for the exact date of returning home from the journey before the most recent journey. The gap time between the two journeys is analysed by non-parametric survival modelling. The number of annual journeys per inhabitant is estimated and a multiplier for each journey is developed. It is shown that the methodology more than doubles the number of reported journeys compared to a traditional retrospective survey reporting the journeys during 3 months

Keywords: Long distance travel; Latest journey; Survey methodology

1. Background

During the last 25 years only relatively few nationwide long distance travel surveys (LDS) have been conducted in Europe and even less have been analysed and presented to the public, especially in journals. In Europe, two main EU research projects have developed a methodology to collect data. The first project, MEST (Methods for European Surveys and Travel Behaviour) from 1996-99 (Axhausen et al. 2003) resulted in the development of the Cross European data collection, Dateline in 2001 which was conducted by the European National Statistics under the direction of Eurostat. The second project, Kite (a Knowledge Base for Intermodal Passenger Travel in Europe) from 2007-09) had three main purposes, 1) to analyse the results from Dateline (Gomes and Santos 2004; Kuhnimhof and

* Corresponding author. Tel.: +45 21231225.
E-mail address: LCH@Transport.DTU.dk

10.1016/j.trpro.2018.10.024

Armoogum 2007), 2) to develop new guidelines for a LDS, (Kuhnimhof et. al. 2009) and 3) to test the method on three surveys from the Czech Republic, Portugal and Switzerland (Frei, et.al., 2010).

Since then, the most recognised European LDS project has been the British survey in which travel behaviour is analysed, and a forecast of long distance travel in UK is developed (Dargay and Clark 2012). Several of the countries conducting a National Travel Survey combine these with a LDS. However, most outcomes are only reported in national reports, if at all, see e.g. for Norway (Vågane, Brechan, and Hjorthol 2011). In Denmark, a LDS was also conducted together with the NTS, but data were never analysed and for economic reasons the survey was stopped in 2001. In 2010-11 a new Danish one-year LDS (only including journeys with overnight stay(s)) was conducted independently of the NTS (Christensen and Knudsen 2015; Knudsen 2015). The reason for separating the surveys was that a very long interview time by telephone was assessed to lead to too many drop-outs if a LDS was included in the daily National Travel Survey. After this survey, only very little interest has been paid to long distance travel and it has not been possible to raise money for a new Danish data collection.

The aim of the present study is therefore to develop a methodology in which the response burden, and thereby the cost, is reduced.

1.1. Traditional Long Distance Surveys

The two most common ways to set up a LDS today are a retrospective survey and a longitudinal survey. In the retrospective survey, the respondents are asked to report all trips they have performed during the latest period, which typically covers between 8 weeks and one year (the so-called travel period). The longitudinal survey follows the respondents who have to report when they are active or at certain intervals, which makes intermediate follow-ups by the organisers possible. According to (Richardson and Seethaler 1999), the two main problems with the retrospective method are:

- A long response period covered by a survey results in a substantial memory effect
- A short response period results in a high share of respondents who have not made a journey. Therefore, the sample has to be rather big to get enough journeys to analyse

The problems with memory effect depend on the frequency of the activity, the length of the trip and if it is private or business (Armoogum and Madre 2002). (Denstadli and Lian 1998) reported a memory effect of 32% for journeys over 100 km whereas longer journeys had a lower memory effect. (Christensen and Knudsen 2015) only find an effect of 16% for journeys with overnight stays. For both, the level is measured as the number of journeys made two and three months back in time compared with the number of journeys during the latest month.

A longitudinal survey, which has to run for a year to take into account the seasonality of long distance travel, solves the problems with the memory effect (Aultman-Hall et.al., 2015). On the other hand, a longitudinal survey suffers from problems with drop-outs during the survey period and might also suffer from some underreporting from those who do not drop out. If each participant is only participating for a short period and is replaced by a new respondent after 2-3 months, the longitudinal study also suffer from the problem of recruiting many participants with only few journeys each.

The interest of the present paper has been focused on a survey based on an idea put forward by (Richardson and Seethaler 1999) who suggested to only ask the respondents to report the latest journey. The results from this should be used to estimate the annual travel frequency and the distribution of trips over the population and types of journeys. As far as the author is aware, only Belgium has used this way of collecting data in practise. However, the author has not been able to find a description of how the Belgian data are collected and handled afterwards.

By using the latest journey methodology, (Richardson and Seethaler 1999) emphasise that the respondents only have to report one journey each. They expect this to increase the number of reported trips to - in principle - one trip per respondent. Furthermore, they consider the memory effect to be smaller than in a traditional LDS and close to zero. Respondents with a low travel frequency are expected to be able to remember their only journey during a long period, and respondents with a high trip frequency will only need to remember a journey that took place a short time ago. In a retrospective survey, high frequent travellers are expected to have low recall of trips during the recall period.

In the following we present a Danish test survey that uses the suggested methodology. A methodology to analyse the data is described and tested on the collected data.

2. Methodology

2.1. The new Danish LDS

The new Danish survey is developed to test the method of only asking the respondents to report the latest journey. As with the survey in 2010-11, long distance travel is defined as journeys with overnight stay(s). Since the Danish NTS (DTU Transport's Data and Model Center n.d.) is a continuous survey with at least 10,000 annual interviews, the responsible staff assess the daily survey to cover domestic long distance journeys independently of length and duration acceptably. The annual kilometres by car is e.g. bench marked against the official kilometres from different calculations and represents 98% of these. International trips are included in the NTS too. However, the number of reported journeys is very low and the details about the international part of the journey is limited to information about the destination. The author therefore assess the main need is to improve the knowledge of the international journeys.

A web survey (CAWI) was therefore set up with the goal of collecting 3,000 interviews in which the respondents were asked to report their latest outbound journey with overnight stay(s). The survey was conducted over a 3-month period during the summer 2017 based on a representative sample of Danish adults aged 16 and over. The respondents were contacted by an e-mail in their electronic mailbox (a so-called e-box). It is mandatory for the Danish population to have an e-box to which public authorities can send e-mails with personal information. However, people without internet access are granted an exemption. 7.4% of the sample has no e-box, the share is increasing by age. A small sample of 600 telephone interviews was collected from non-respondents (a third from the group without e-box) to uncover the size of a possible bias. Only the results from the CAWI are included in this paper.

The sample was evenly distributed over each day during a 12 weeks' period starting on June 23, i.e. 239 persons were allocated to each day. The letters were sent in batches once a week, normally on Thursdays, and received by the respondents the same day. After two weeks, a reminder was sent to those who had not participated yet. An important problem using the e-box solution is that many people only open their e-box if or when they get a letter they consider relevant for them. Due to the data collection period taking place during the summer, the results are also influenced by seasonality. 5 weeks after the first e-mail we realized that the response rate was too low to reach the intended 3,000 interviews. It was decided to send a reminder by SMS to people with a known cell phone number. The last reminder was sent in the second part of September and data collection was closed by the end of October when only one daily response was received in the database. The final response rate is only a bit over 15%.

The main part of the questionnaire is allocated to the latest journey. The respondents were asked in which year the journey took place. In case it took place in the period 2015-17 they were asked to report many details, including the exact date of returning home (if not available, the week or month). For earlier journeys, fewer details were requested. To support the development of the methodology of travel frequency, the respondents were furthermore asked to report two more journeys with a private purpose in the period 2016-17. For these, only the main information about the journeys was collected on top of the exact date. Other subjects in the survey were background information, including availability of holiday homes, number of international business and study journeys during the latest year and information about an international one-day trip during the latest month. Finally, questions about attitudes to use of different modes, motives for holiday travel and environmental norms were included.

The length of the CAWI interview resulted in 9% of the participants dropping out after they had started to answer the questions about the latest journey. Respondents who have provided information relevant for analyses are included, even if the interview is interrupted at a later stage.

2.2. Analysis of the latest journey

A main purpose of the project is to develop a methodology to determine the travel frequency of the respondents measured as number of trips per year. In a normal LDS covering a response period of e.g. 3 months, the number of annual journeys is derived by multiplying with 4 the number of reported journeys during the 3-months period, while ignoring the memory effect. A survey including only the latest journey is missing journeys performed by the respondents who have more than one journey during the year. It is therefore necessary to use knowledge about the reported journeys to infer the missing journeys.

Furthermore, the results are biased if travel distances and travel behaviour is analysed based on the pool of all journeys in the survey. Journeys taking place a year or more ago are probably not representative for the journeys taking place shortly before the survey. It is needed to develop a probabilistic trip rate multiplier as a function of the number of days since the journey was made. This rate has to be used as a weight for analysing the distribution of mode choice, travel purpose etc. It is especially necessary to up-weight travel distances for the journeys to calculate the correct annual travel distance per inhabitant or per mode.

Based on simulations, (Richardson and Seethaler 1999) have developed a probabilistic model for such an average trip rate multiplier as a function of the number of days since the journey was made to calculate the full number of annual journeys. Their method is developed under ideal conditions such as:

- no seasonality
- journeys can last only one day
- constant travel frequency per traveller
- no memory effect
- a simple assumption of the distribution of the population's travel frequency as linear

They find that the travel frequency of the traveller returning home on the survey day has to be multiplied by 1.61 because travellers travelling the day after the former journey have a high probability to travel very often. The multiplier is decreasing the further back in time the latest journey was made. A journey made around 100 days ago has to be included by a weight of one whereas a journey made a year ago only has to be included with a weight of 0.4 because it represents few journeys. By simulation they show that for a 30-day reporting period, the application of a probabilistic multiplier to the most recent journey produces trip rates close to those obtained from an ideal 3-month retrospective survey with respect to both average and variability.

When collecting real data, a model developed under the described ideal conditions is not feasible. First of all the assumption about the distribution of number of journeys over the population is unknown, because it cannot be revealed from a three-month survey. In the Danish survey from 2010-11 nearly half of the respondents is not travelling and their distribution of travel frequency is unknown. 10% of the respondents represent 50% of the journeys which is not indicating the chosen linear distribution. Another problem is the assumption about journeys lasting one day when only journeys with overnight stays are included.

2.3. Survival modelling

Instead we turn to the method suggested by (Frei et al. 2010). They conducted a long-distance travel survey covering 8 weeks. However, if the respondent made no journeys during the 8 weeks they asked about the date of a possible journey after the beginning of the 8 weeks or the date of the latest journey before the first journey in the 8-week period. The added journeys are included as censored journeys with no other information than the date (in fact, this author would have added the journeys as real incidents). The time from the survey date to the latest journey and the time between the reported journeys were modelled by survival and hazard modelling using a non-parametric survival distribution. From this they estimated the hazard ratio which was used to up-weight the annual number of journeys for the surveyed respondents.

Survival analysis and hazard modelling is developed to estimate e.g. survival time of a population or a group of people suffering from a disease, to test the effect of medical treatment in epidemiology and pharmacy or to analyse breakdown of technical products. In transport research, survival modelling has mainly been used for analysing duration of transport activities, e.g. (Bhat 1996).

In epidemiology, the time to an event is normally censored meaning that it is not the whole sample that has experienced the incident (yet). Some of the test persons might die from other reasons and have therefore left the investigation. Others have not experienced the analysed incident (yet) when the experiment / investigation ends. The same is the case with our travel data between others because many respondents have not taken part in two private journeys after January 1st 2016 at which a limit for travelling was set up. (Lu and Shen 2014) emphasise that special methods are needed when handling censored data because ignoring the censored observations and using traditional multivariate regression methods will result in serious mistakes when the censored observations reach a certain amount.

2.4. Gap times and censored gap times

The analyses only include private journeys because the two journeys before the latest journey are only private. Therefore, business journeys are removed. The analysed time is mentioned by the term 'gap time'. The first gap time starts on the date allocated to the respondent and on which the journey had to be finished at the latest. It ends on the date the respondent returns home. The second and third gap times are the time gap between the latest journey and the latest private journey and between the two private journeys, respectively. These gap times starts when the respondent returns home from one journey and ends when he/she is back from the next journey. The first gap time can be zero. However, to be able to use the logged value (which is used in the estimation) it is offset with one in case of the value one.

The latest journey may be reported on any date back in time, whereas the two private journeys are only reported if they took place in 2016 or 2017. Therefore, in many cases only one earlier journey is included and sometimes none at all. When the investigation period ends before a private journey took place, a censored 'journey' is added with a gap time as the time between the latest reported journey and the earliest date it could take place i.e. January 1st 2016. In case the latest journey is a business journey or the reported latest journey by a mistake end after the stated limit for returning home a censored gap time is added as the third or second gap time. Furthermore, some journeys have been removed because they are obviously the same journey as the former one. The gap time is replaced by a censored time, too. In these cases, the censored gap time is set to the day before the outbound trip of the first known journey.

Sometimes the respondent declares that he/she had not been travelling at all before the already reported journey(s). In this case, the censored gap time should go back to the respondent's birth year. The same is the case for the few respondents who have never travelled abroad. However, as the survey only includes persons aged 15 or above it should not be possible to include journeys that have taken place before the end of year when the respondent turned 16. In both cases the censored gap time is dated to end of the year when the respondent turned 16. Similarly, journeys taking place in an early age are removed and replaced by a censored gap time.

2.5. Survival modelling

In survival modelling the cumulative distribution function *F(t)* describes the probability that the random time *T* to an event will be less or equal to a chosen date (e.g. Kalbfleisch and Prentice 2002; Lu and Shen 2014):

$$F(t)=P[T \leq t],\ t \geq 0. \tag{1}$$

The derivative of *F(t)* is called the probability density function *f(t)*. The survival function *S(t)*, which is the probability that an event takes place later than a specified time is defined as:

$$S(t) = P[T > t] = 1 - F(t) = 1 - \int_{x=0}^{x=t} f(x)dx \tag{2}$$

S(t) is a non-increasing function with its values between 1 and 0. When *t* approaches ∞, *S(t)* approaches 0. In our case, t never approach ∞ for any of the three gap times because some of the respondents have not experienced a journey or had not made two, respectively three journeys since January 1st 2016.

For each of the survival functions there is a maximum value (MAX) of t. The survival function therefore has to be corrected by normalisation of the density function. The normalised density function is

$$\tilde{f}(t) = [S(t) - S(t+1)] \mid t < MAX = \frac{f(t)}{\int_{x=0}^{x=MAX} f(x)} \tag{3}$$

The normalised survival time, which is 0 by the maximum gap time is therefore:

$$\tilde{S}(t) = 1 - \int_{x=0}^{x=t} \frac{f(x)}{\int_{y=0}^{y=MAX} f(y)} \quad (4)$$

The mean value μ(t) of the survival function is the area under the survival curve from gap time 0 to gap time t. μ(t) is the estimated gap time at time t.

The survival function is estimated as a Kaplan Meyer estimate and fitted by a non-parametric survival distribution (Kalbfleisch and Prentice 2002). For the estimations is used the SAS procedure Proc Lifetime (Lu and Shen 2014). When using a non-parametric form no assumption is made about the distribution of the gap times which means that the estimation can get 'closer' to the real distribution of the observed gap times. However, it also means that the mean gap time has to be calculated numerically.

In the following is shown how the estimated survival function is used to estimate the annual number of journeys per respondent.

3. Results from estimation

3.1. Estimation of the survival model

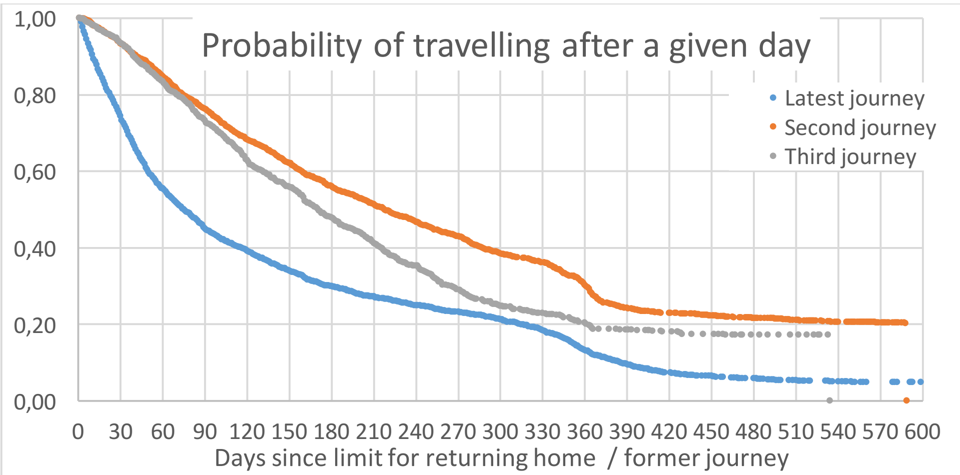

Figure 1 Survival curves showing the probability of ending a journey after a certain date (The normalised survival functions are shown)

Figure 1 shows the estimated survival curves for the three gap times. The survival curve for the first gap lies below the curves for the two other gap times. This means that the probability of travelling earlier than e.g. 3 months is smaller than the probability of travelling again more than three months after the first journey. Both the Rank Test and the Wilcoxon test shows with a p-value less than 0.0001 that the three survival curves are not similar. This is, however, what could be expected. In theory, the time from the date of the survey until the most recent journey is only half of the time between two random journeys when the survey date is chosen independently of the travel dates. This is the same as the theory for the waiting time at a bus stop in case of random arrival of the travellers. It is half of the time between two bus departures in case of fixed frequency.

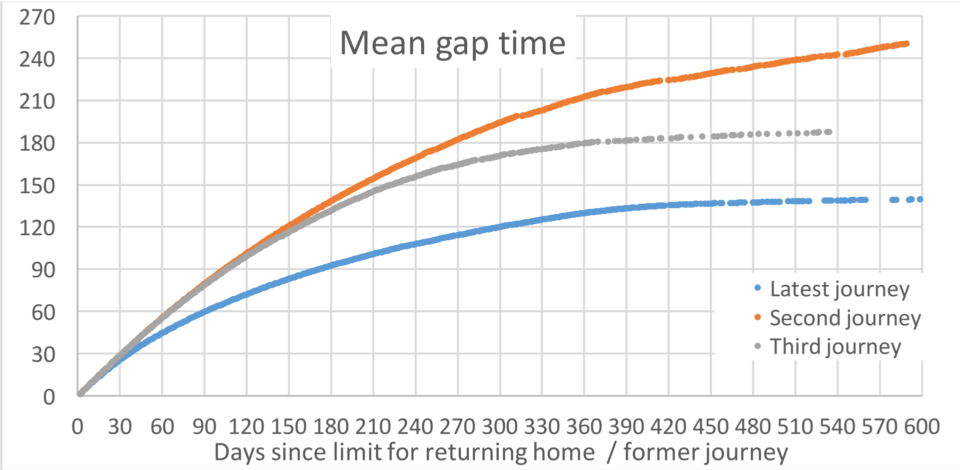

Figure 2 The estimated normalised mean gap time for the three journeys.

The survival curves for the gap time between journey one and journey two and between journey two and journey three should on the other hand be the same when there is no seasonality in the distribution of the journeys, i.e. random travel dates. As can be seen, the survival time for gap time two and gap time three is overlapping for the first six months. After this, the time to journey three is shorter. It is tested if the survival function for the second and third journeys can be accepted as similar, when controlling for the survey week for the respondent (the week with the date for the latest day of the homebound journey). The two functions are still different.

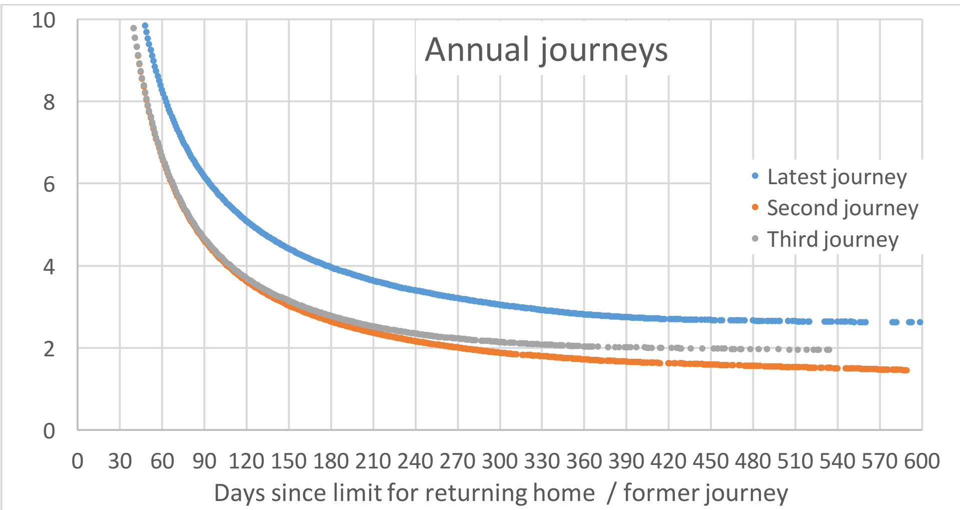

Figure 3 The estimated normalised number of annual journeys.

For all three gap times, a slower decrease in the probability of travelling is observed from around ¾ year after start of the gap period and until around a year after the start. The effect is largest for the first gap time and smallest for the third. This shows a seasonality in data representing the former spring and summer holiday. This seasonality is probably an important reason for the difference between the survival functions for the second and third gap time.

Figure 2 shows the mean gap time between two journeys, again calculated as the normalised mean gap time. The mean gap time is the area under the survival function at the number of days it is calculated for. The mean gap time is used to calculate the annual number of journeys for a given period, see Figure 3. For one year, the annual number of journeys is 1.71 when calculated from the gap time between the first and second journey.

3.2. Comparing results with the Danish long-distance travel survey from 2010-11

In the present study, 3,100 respondents have answered questions about their latest journey. Of these 0.5% had no journey to be included and 6.3% has not reported details about the journey because they only travelled before 2012 (5½ years ago). Only those who travelled during the latest 2½ years were asked to report all details about their latest journey. They represent 88.2% of the respondents. Of these 0.9% stated that they had no earlier journeys than the reported. This means that 87% of the respondents delivered answers useful for the analyses.

The Danish retrospective survey from 2010-11 showed that the population made 1.45 international private journeys and 0.35 business journeys per year. With a surveyed travel period of three months, 29% of the respondents were travelling internationally.

If we make 10,000 interviews using the suggested alternative method and include details about journeys made during the latest 2½ year, we will get information about 8,700 private journeys. If we make 10,000 retrospective interviews covering a 3-month period, we will get information about journeys from 2,900 respondents reporting 3,600 private journeys or less than half of the journeys.

3.3. Memory effect

An important question is the quality of the responses, when collecting information about journeys far back in time. Most important is the memory effect. n the long-distance travel survey from 2010-11 is asked for the number of journeys month by month during the three latest months. The number was divided into domestic and international journeys and into journeys with 5+ nights' stay and less. When the latest month is taken as the bassline, the overall number of reported journeys is 22% lower one month earlier and 19% lower two months earlier. When only looking at international journeys the same figures are 18% and 16% respectively. Furthermore, the respondents are asked how many journeys they have had with 5+ nights in the nine month earlier (up to one year overall). With the same methodology, the result is 43% lower for international journeys and 39% lower for domestic journeys.

A distribution of the journeys over the year of the baseline month shows clear peaks in April (Easter), May (several short holidays) and July (summer holiday). The peaks for month two is less distinct and in the third month, the peaks have nearly disappeared. Only a small peak is seen for the summer. This indicates not only a memory effect at which journeys are forgotten but probably also a memory effect at which earlier journeys are reported as conducted in the 3-month period. The travel frequency is therefore debatable.

Unfortunately, it is not possible with our tree months' survey to compare the memory effect by comparing the latest months with the former months when the interviews only have been collected during three months. With a survey running over a year, it would be possible to compare the distribution over a year of the latest journeys dependent on the survey month. Still, due to the substantial seasonality, it will not be possible to get a precise answer to the recall effect.

4. Conclusion and recommendations

The aim of this study is to develop a survey methodology for collecting travel data for long distance travel, which can increase the amount of reliable data about the journeys at a reduced cost.

Only 29% of the respondents in a Danish retrospective survey in 2010-11 have been travelling internationally during the last three months. Some of these have never been travelling; others might have travelled 4 months ago. This makes a substantial difference. On the other hand, 10% of the respondents have been travelling at least 10 times during the 3–month survey period (both domestic and international journeys). They represent 50% of the journeys. Together, this might bias the understanding of the resulting travel pattern, and especially of the behaviour of travellers. Furthermore, the high frequent travellers easily forget a journey or two, which is documented by a bias in the number of journeys one, two and three months back in time. Due to the high share of respondents without any journey, it is needed to collect many interviews to get a reliable number of journeys to analyse.

Instead, we as an outset proposed to ask for the latest journey only. However, the analysis shows that it is not a correct solution to ask for the most recent journey only. It is necessary to ask about the date of the latest journey before that too. With this information, it is possible to use a non-parametric survival modelling procedure to estimate a survival function for the gap time between the most recent journey and the latest journey before that. This survival function is normalised due to the cut off at the highest observed gap time. Based on the survival function for the second journey, the mean gap time and the annual number of journeys can be estimated.

It has been shown that the two-latest journeys methodology will produce information about more than twice as many journeys as a retrospective survey reporting journeys in a 3–month period. In the Danish case, 87% will deliver information about their international journey with the new survey whereas only 29% deliver data in the old way. This means that the number of interviews can be more than halved to get the same precision. The average interview time will increase but less than the time saved by reducing the number of interviews.

What we do not know is to what extent the new survey is suffering from memory recall regarding details about the two most recent journey. Especially, mistakes in the travel dates will bias the results. This can only be analysed when a survey covering one full year is available, and even then only with difficulty. Another weakness is that it is not possible to point out the high frequent travellers and analyse their travel behaviour separately. Off curse, travellers with 2 or 3 journeys during a short period are more frequent travellers that those who have not been travelling during one year. However, the 2-3 journeys can be accidental and the gap to a former journey can be high. Asking about the behaviour through 3 months is providing more information about the high frequent travellers but less about the low frequent travellers.

It should be mentioned, that the two methodologies can be combined similar to what was done by (Frei et al. 2010). You can ask for the two latest journeys and supplement with a number of journeys performed during the last 8 weeks or three months. In our survey we asked about the number of business journeys during the recent year. This can also be used to identify the high frequent business travellers, even though the exact number of journeys is uncertain. Similarly can be done for holiday journeys and visits to family and friends.

For the moment, all journeys in our survey are weighted equally. Further theoretical development has to be done to be able to up-weight the higher frequent travellers' journeys to be able to calculate a correct mode share, share of purpose and especially the annual travel distance. If the annual travel frequency is used as weights, journeys conducted short before the former has a very high weight. (Richardson and Seethaler 1999) only finds a small overweight of the journeys short before the latest. With our method, such problems cannot be avoided. However, the effect is reduced when data is collected during one year because the summer holiday generates more journeys with a short gap time. The effect can be studied more deeply with a full year survey.

4.1. Recommendations for details in a new survey

It should be mentioned that the method used in the present survey to ask about journeys undertaken in a limited period should not be repeated because it reduces the mean gap time for the journeys taking place close to the beginning of the period. Instead, the date should be open-ended. In case of fear of a memory recall effect or drop-outs due to respondents' problems with answering the question precisely, it is better to accept an imprecise answer, for instance only the month for journeys made more than one year back in time, because the effect on the mean gap time is limited. An imprecise answer is probably better than a censored gap time when estimating the survival model. The need to normalise the survival function also increases the uncertainty of the results. The closer the survival function gets to zero by observed journeys the more correct the mean gap time can be estimated (the area under the survival curve)

On the other hand, it is important to get the date precise for journeys, which were finished short before the former journey. Mistakes in these journeys lead to a change in the model and in the annual number of journeys. This is an important weakness of the proposed method. You should as a control ask for the duration of the latest journey and if the respondent can remember how many days they stayed at home before the next journey started. This is especially important if the respondent reports two journeys in the same month or week without specifying the exact dates.

It is furthermore important to refer to the date of returning home and not to the start of the journeys, to be sure to include all journeys finished at the allocated date. Using the date of start of the journeys would result in long-stay journeys being left out. Another important issue is, that the latest date of finalising the latest journey has to be fixed and evenly distributed over the year. The date should not be up to negotiation and journeys which are finalised too late, have to be removed. A warning to the respondent should be included already during the interview.

In our survey is asked for many details about the most recent journey and only the main questions are asked about the earlier journeys. It might be more correct to ask for details for the journey before the most recent because it is more random. On the other hand, asking about an earlier journey instead of the most recent will result in more memory effect. Furthermore, the further back in time the journey took place, the higher is the probability for a change in income, family situation and location of the residence. The gain from a theoretically more random answer is therefore weighted out by a lower quality.

References

Armoogum, Jimmy and Jean-Loup Madre. 2002. "Weighting and Correcting Long-Distance Travel Surveys." Pp. 151–69 in *Capturing Long-Distance Travel*, edited by K. W. Axhausen, J.-L. Madre, J. W. Polak, and P. L. Toint. Research Studies Press, Baldock.

Aultman-Hall, Lisa, Chester Harvey, Jeffrey J. LaMondia, and Chloe Ritter. 2015. "Design and Response Quality in a One-Year Longitudinal Survey of Overnight and Long-Distance Travel." *Transportation Research Procedia* 11:136–53. Retrieved (http://dx.doi.org/10.1016/j.trpro.2015.12.012).

Axhausen, K. W., Jean-Loup Madre, John W. Polak, and Ph. L. Toint, eds. 2003. *Capturing Long-Distance Travel*. Baldock: Research Studies Press LTD.

Bhat, Chandra R. 1996. "A Hazard-Based Duration Model of Shopping Activity with Nonparametric Baseline Specification and Nonparametric Control for Unobserved Heterogeneity." *Transportation Research Part B: Methodological* 30(3 PART B):189–207.

Christensen, Linda and Mette Aagaard Knudsen. 2017. *Long Distance Travel - A Study of Dane's Journeys during 15 Years*. Lyngby: DTU Transport Report 10.

Dargay, Joyce M. and Stephen Clark. 2012. "The Determinants of Long Distance Travel in Great Britain." *Transportation Research Part A: Policy and Practice* 46(3):576–87. Retrieved March 17, 2014 (http://www.sciencedirect.com/science/article/pii/S0965856411001856).

Denstadli, Jon Martin and Jon Inge Lian. 1998. *Memory Effects in Long Distance Travel Surveys*.

DTU Transport's Data and Model Center. n.d. "The Danish National Travel Survey, TU." *Data 1975, 1981, 1992-2015*. Retrieved November 6, 2017 (http://www.modelcenter.transport.dtu.dk/english/TU).

Frei, Andreas, Tobias Kuhnimhof, and Kay W. Axhausen. 2010. "Long-Distance Travel in Europe Today: Experiences with a New Survey." P. 16 in *Transportation Research Board*. ETH, Eidgenössische Technische Hochschule Zürich, IVT, Institut für Verkehrsplanung und Transportsysteme. Retrieved July 10, 2014 (http://trid.trb.org/view.aspx?id=910298).

Gomes, Faustino and Fátima Santos. 2004. "Design and Application of a Travel Survey for Long-Distance Trips Based on an International Network of Expertise." P. 11 in *European Transport Conference*. Association for European Transport.

Kalbfleisch, J. and R. Prentice. 2002. *The Statistical Analysis of Failure Time Data*. Wiley-Interscience. John Wiley & Sons Inc, Hoboken.

Knudsen, Mette Aagaard. 2015. "Danish Long Distance Travel - A Study of Danish Travel Behaviour and the Role of Infrequent Travel Activities." Technical University of Denmark.

Kuhnimhof, Tobias, Roger Collet, Jimmy Armoogum, and Jean-Loup Madre. 2009. "Generating Internationally Comparable Figures on Long-Distance Travel for Europe." *Transportation Research Record: Journal of the Transportation Research Board* 2105:18–27. Retrieved August 8, 2013 (http://trb.metapress.com/openurl.asp?genre=article&id=doi:10.3141/2105-03).

Kuhnimhof, Tobias and Jimmy Armoogum. 2007. "Existence and Comparability of Data Sources." 65. Retrieved (http://kite-project.eu/kite/cms/images/stories/kite/WP2/D3 - Existence and Comparability of Data Sources.pdf).

Lu, Jane and David Shen. 2014. "Survival Analysis Approaches and New Developments Using SAS." P. 13 in *PharmaSUG*, *PharmaSUG*. Retrieved (https://www.pharmasug.org/proceedings/2014/PO/PharmaSUG-2014-PO02.pdf).

Richardson, A. J. and R. K. Seethaler. 1999. "Estimating Long-Distance Travel Behavior from the Most Recent Trip." P. 18 in *Personal Travel:The Long and Short of It*, edited by TRB Committee on Travel Survey Methods (A1D10) and TRB Committee on National Transportation Data Requirements and Programs (A5016). TRB Transportation Research Circular E-C026. Retrieved November 8, 2016 (http://ntl.bts.gov/lib/9000/9200/9200/09_richardson.pdf).

Vågane, Liva, Inge Brechan, and Randi J. Hjorthol. 2011. *Den Nasjonale Reisevaneundersøkelsen 2009 - Nøkelrapport*. TØI rappor. Oslo: Transportøkonomisk Insttut. Retrieved (https://www.toi.no/getfile.php/Publikasjoner/TØI rapporter/2011/1130-2011/1130-2011-el.pdf).

Acknowledgement

The author want to thank Professor Jeppe Rich and associate professor Anders Stockmarr for support with developing the methodology for analysing the collected data.