Travel-time ratios for commuters in the Netherlands between 2005 and 2009.

June 2013

Dennis Hunink
Student at the University of Utrecht,
Department of Social Geography.

Table of contents

Introduction

For every hour that commuters work, they travel a single trip of 3,5 minutes.
This statement comes from an empirically study on the Dutch National Travel Survey 1998 data by Schwanen & Dijst (2002, p. 579). With their study Schwanen & Dijst contributed to a growing interest (Xuedong & Pas, 1999; Dijst & Vidakovich, 2000; Schwanen & Dijst, 2002; Chen & Mokhtarian, 2006) in the relationship between participating in activities and the need to travel that derives from that participation (Recker et.al, 1986, p. 310).

This empirical evidence was based on the theoretical concept of the 'travel-time ratio', a concept that has been published and empirically tested in a 2000 article by Dijst & Vidakovich. The empirical findings, mostly provided with the concept itself, of both the 2000 and 2002 studies have been an anchor for further research in transportation studies.

But a lot has changed in the world since the 1998 National Travel Survey data has been collected. Furthermore, Mokhtarian & Chen (2004) conclude from a review of various studies that it's hard to state that travel times stay the same over time.
As travel time is not constant, it is possible that the travel times that have been found by Schwanen & Dijst (2002) could have changed over time. For example, the use of Internet and Communication Technologies has been found to have an impact on people's travel times (Lyons & Urry ,2005; Wang & Law, 2007).
Therefore the present article will try to give an updated travel-time ratio on the aggregate level for commuters in the Netherlands by using more recent data of the Dutch National travel Survey. Furthermore, this article aims to provide more details about the methodological choices in selecting data and the way different selections influence the results of analysis.

Taking the aggregated results in the article of Schwanen& Dijst (2002) as a starting point of interest, this paper raises and tries to answer the question whether similar results can be derived from more recent data, using the travel-time ratio concept, accompanied with a different way of data selection. The answer to this question may be relevant for scientists that are dealing with travel-time ratios, both when reading previous studies as when reviewing upcoming research.
It may also in some way be of benefit to policy makers, who are dealing with travel behaviour of commuters and are trying to develop ways to influence the time commuters spend on the road, or even for other professionals that have to deal with travel behaviour and in particularly travel times for their profession.

Theory

In the analysis on the aggregate level as conducted by Schwanen & Dijst (2002) they find that the travel-time ratios of Dutch commuters can be divided in three groups. The first group consists of people who work up to four hours a day. It turns out their commuting time tends to be stable and therefore not depending on the duration of there stay on the workplace. The second group of commuters shows a different result. For people in this group, who work up to eight hours a day, the commuting time tend to rise similar with the duration of workplace stay. Finally, the people who work more then eight hours a day, again have a stable commuting time. This could possibly be explained by biological factors; after eight hours of working only a certain amount of energy is left to spend on commuting.

The journey of a commuter could consist out of multiple trips (Susilo & Dijst, 2009, p. 21), for example when one has a different working location in the morning then in the afternoon or if one stops at the gym for an hour on his way back home. To distinguish the differences in travel patterns and travel-time ratio's for journeys consisting out of more then two trips, very detailed information is needed to, for example, perform GIS-analysis. However, this paper doesn't seek to find differences between different types of commuters. Instead, it tries to compare the travel-time ratio of one type of commuters throughout multiple years. Therefore the group of commuters to be analysed has been defined as the so-called classic commuters, the commuters who are traveling from home to work and back on the same day, without traveling to other locations to participate in a separate activity on the same day, as shown in Figure 1. In literature about travel behaviour this is referred as a trip with two base locations and no additional locations at all (see for example Susilo & Dijst, p.4).

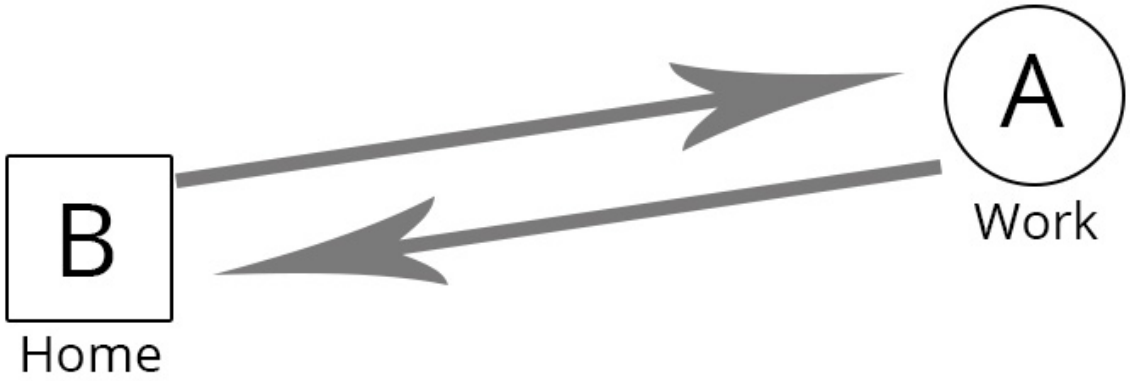

Figure 1. Visual representation of a classic commuting round trip

To analyse the relationship between work duration and commuting time Dijst & Vidakovic (2000) developed the concept of travel-time ratio. This travel-time ratio is obtained by dividing the travel time by the sum of the travel time and activity duration. It can be wrapped in a formula as followed, where T_t represents travel time, T_a activity duration and τ the actual travel-time ratio.

$$\tau = \frac{T_t}{T_t + T_a}$$

Prior to actually calculating a travel-time ratio, data must be selected. The data selection of Schwanen & Dijst (2002) is based on selecting the 90% most probable observations, which leaves the 5% longest and 5% shortest visits to workplaces out of their analysis (Schwanen & Dijst, 2002, p. 577). However, this statistical selection method doesn't exclude one-way trips. Since the concept of travel time ratio is based upon comparing round trips to activity duration, only the data of round trips should be included in the analysis to get a more accurate result.
The present article aims to use a different approach on selecting the data by testing every data entry to a couple of conditions. With this technique it is possible to determine which records are part of a round trip and more important, which records are not. One of the goals of using a 90% statistical selection is to stress out the effect of falsely administrated records (Schwanen & Dijst, 2002, p. 577). With the approach in this article, these cases are already eliminated.

Research Design

This chapter describes what kind of data has been used for the analyses, how the data selection has taken place and how the outcome of the data selection will be analysed.

Characterises of the dataset

The data used for the analysis has its origin in the Dutch National Travel Survey, NTS for short. The travel survey is carried out yearly by Statistics Netherlands in cooperation with the Dutch ministry of Infrastructure and the Environment. The survey has had a couple of different names and was carried out by different governmental organisations, since it started in 1985. In the analysis presented here, the data is provided under the name National Travel Survey.

The research process of the NTS consists of four steps. In the first step a sample is derived out of the entire Dutch population. After sampling the data is actually being collected by surveys and interviews. Thirdly the collected surveys and interviews are processed and the dataset is build. Finally the data in the NTS dataset is corrected and weighted so it can give the most representative perspective for the Dutch population (Ministerie van Verkeer en Waterstaat/Rijkswaterstaat, 2009, p.12).

Each of the datasets of the different years contains over a hundred thousand movements. Each movement is stored as a record and each record has more than a hundred variables, varying from mode of transport to start and end location. A couple of the variables contain details of the person who committed the movement, such as household income, age categories and employment status. For each record there are factors available which makes it possible to weight records and draw conclusions based upon the dataset that are eligible for the complete Dutch population.

The conceptual data selection method

In order to draw conclusions about the travel-time ratios of commuters in this paper, two things have to be done. The first is making the appropriated selection; the second is the calculation of the travel-time ratios for the selected cases.

Selecting the commuters that are the subject of the analysis in this paper requires a very strict definition that is beyond any reasonable doubt. Therefore only classic commuters will be part of the selection. Selecting them has to be done using multiple steps. A first, rough selection can be made by selecting only those records of the NTS dataset that have 'from or to work' as travel motive. See Table 1 for the number of records selected. At this point, no additional criteria have been used.

Year	Number of households	Number of persons	Number of movements	Number of movements with motive 'from or to work'	
				Absolute	Relative
2005	28.436	64.052	196.075	38.624	19,7%
2006	23.695	53.545	165.521	32.246	19,5%
2007	23.240	52.218	159.637	30.347	19,0%
2008	18.102	40.125	121.107	25.761	21,3%
2009	18.158	40.836	123.870	25.247	20,4%

Table 1. Number of records per year

The second step of the selection consists not so much of reviewing values of variables of single data records, than of focussing on comparing multiple records with each other. This has been done by using custom written PHP scripts that include multiple SQL queries. Both PHP and SQL are programming languages used by web- and application developers, which turned out to provide a flexible and accurate solution to compare data records in a dynamic way. What the scripts have done is select all and only the records that meet all of the following criteria, as visualized in more detail in Figure 2:

1. The person, who travelled, only travelled twice on that specific day.
2. The postal code of arrival of the one trip is the postal code of departure of the second trip.

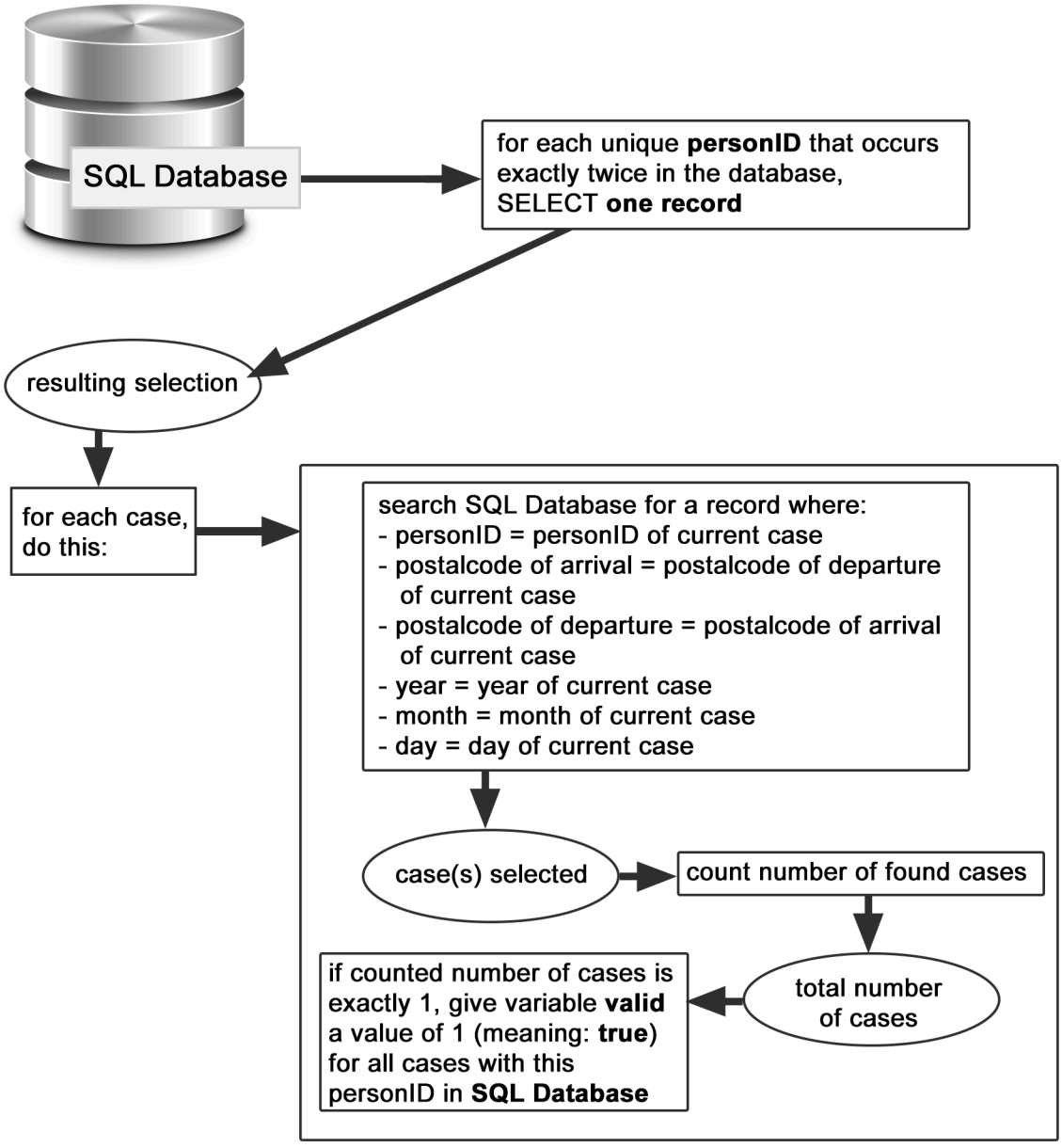

Figure 2. Visual representation of PHP data selection

It turns out that from the amount of trips that was selected at first, because they had 'from or to work' as a motive of travelling, only slightly more then half of them passed this second selection, as shown in Table 2.

Year	Trips total	Valid trips		Invalid trips		Round trips
		Absolute	Relative	Absolute	Relative	
2005	38624	21244	55%	17380	45%	10622
2006	32462	17284	53,2%	15178	46,8%	13225
2007	30347	16606	54,7%	13741	42,3%	8303
2008	25761	14120	54,8%	11641	45,2%	7060
2009	25247	13862	54,9%	11385	45,1%	6931

Table 2. Number of selected records per year

After creating the selection, the actual travel-time ratio can be calculated. To do so, the formula that was constructed by Dijst & Vidakovich when they introduced the travel-time concept (2000, p. 187) is used:

$$\tau = \frac{T_t}{T_t + T_a}$$

$$\tau = travel\ time\ ratio$$
$$T_{=t} = travel\ time$$
$$T_a = activity\ time$$

One could choose to simply double the travel time for each record and then calculate the travel time ratio, as has been done by Schwanen & Dijst (2002, p. 577). Although differences in outcome tend to be rather small, they do differ from exactly filling in the travel time formula as it is strictly written down. That do is possible within the PHP and SQL scripts mentioned before, since the activity time is know for each single record and the records that belong to the same journey can be linked to each other, the travel time-ratio can be calculated when the PHP/SQL finds that a record meets the criteria mentioned before. The following example demonstrates how the formula can be used differently and produces different results, although the difference is rather small.

Example: two trips that form one journey.
Trip 1: travel time = 40 minutes, activity time is five hours.
Trip 2: travel time = 90 minutes, activity time is also five hours, since belonging to the same journey.

Calculated by doubling trip time (Schwanen & Dijst, 2002 method):

$$\frac{40 * 2}{40 * 2 + 300} = 0{,}211$$

$$\frac{90 * 2}{90 * 2 + 300} = 0{,}375$$

Travel time ratio for this journey: 0,211+0,375 / 2 = 0,293

Calculated for journey as whole (original method):

$$\frac{40 + 90}{40 + 90 + 300} = 0{,}295$$

Travel time ratio for this journey: 0,295
Difference between the two methods is 0,002.

After calculating the travel time-ratios, the data is analysed using SPSS statistics. Because the nature of the data, it is possible to generalize the results too the level of the entire Dutch population. To do so, there are three variables available for each record: one to weight for the person conducted the trip, one for the household he/she belongs to and one to weight the trip itself. Since the analysis focus on the trip itself, the last variable has been used.

General results

In this section the results of the travel-time analysis are drawn. The main focus of this chapter is the influence of the new selection technique on the conclusions resulting from the graphs drawn.

The travel-time ratio in The Netherlands

The mean travel-time ratios vary from 0,189 to 0,206 in the period between 2005 and 2009, as Figure 3 shows. Although the differences may seem small, they are statistically significant. Both 2006 and 2009 stand out, when tested with Anova. The first for being significantly higher than other years, the latter for being most equal to the other years. Since this has been tested with an Anova test and one year stands out, namely 2006, all other years will have a bigger chance of being significantly different as well. Because the number of cases is very large, the latter cannot be completely accounted to the high mean travel-time ratio of 2006 but should also be addressed to the records of the years themselves.

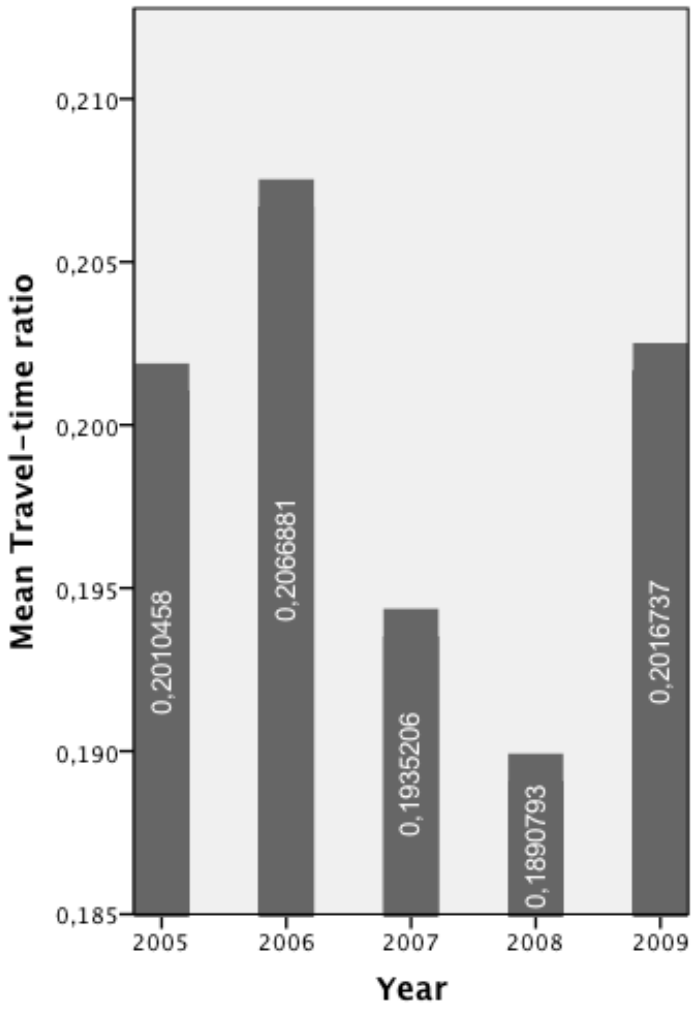

Figure 3. Mean travel-time ratio between 2005 and 2009

To review the results in more detail, two years will be investigated in more depth. To illustrate the effect of the selection technique used in this paper, 2005 and 2009 suit

best, being the first and last year of all years reviewed. Though only these two years will be drawn in detail here, the other years do not show very different results.

Starting with 2009, the scatter plot in Figure 4 shows there is a concentration of cases for a work duration of 1 up to 630 minutes. The figure also reveals that the number of cases decreases when work duration goes up. As it is common practice in previous studies, Schwanen & Dijst (2002, p.577) took only the 90% most probable cases into account, which would especially leave the cases with a the highest work duration out. By doing so, they where able to draw a line graph of the mean travel-time ration in which three different phases can be distinguished: up to 240 minutes, 240 to 480 minutes and more than 480 minutes of work duration (Schwanen & Dijst, 2002, p.581).

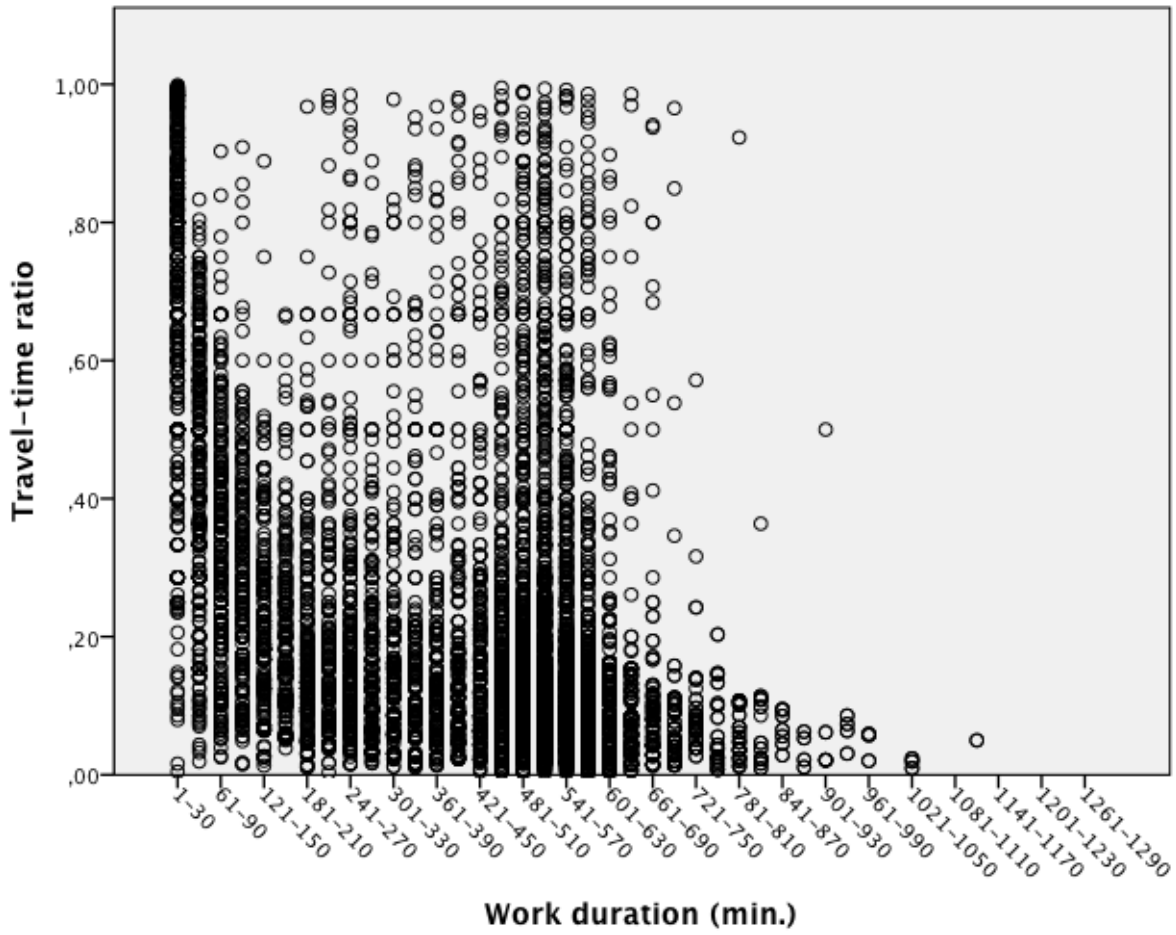

Figure 4. Scatter plot for all cases in 2009

Similar to Schwanen & Dijst (2002, p.581), a graph has been drawn in Figure 5, consisting of only the 90% most probable observations in 2009. The two vertical non-dashed lines represent the boundaries of the three phases found in 2002, as mentioned above. Alternatively, two dashed lines represent new boundaries for 2009. They are placed on points where both the mean travel-time ratio and the mean activity duration show a trend that differs from the trend in the phase before.
On the basis of this graphic, one could argue whether or not it is valid to divide the graph in three stages. In the context of this paper, one remark on that subject has to be made.

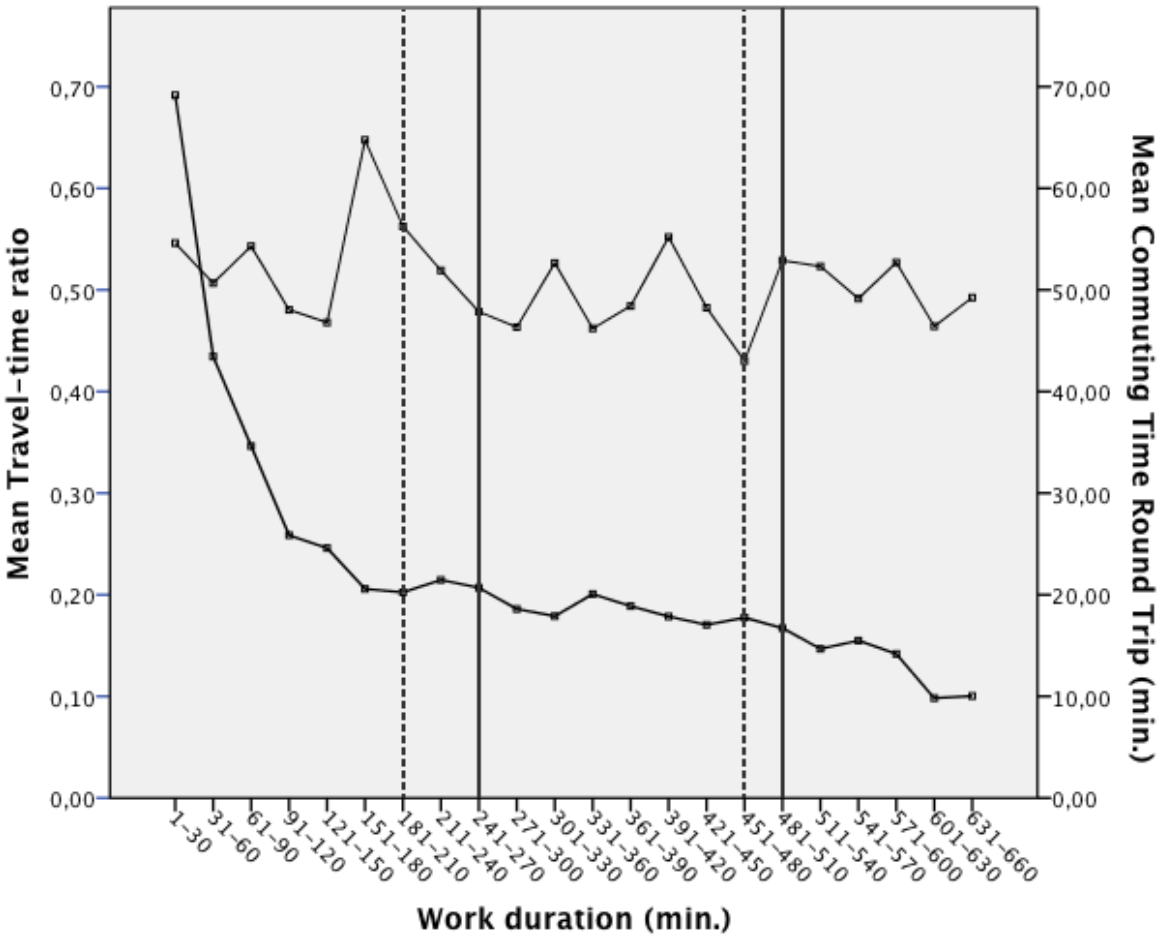

Figure 5. Travel-time ratio and work duration for 90% most probable cases in 2009

The foundation of the graphic lies in the selection of the 90% most probable cases, to avoid for instance wrong administrated cases to be part of the analysis. As explained earlier, this paper uses a alternative selection method to make sure that only valid cases of classic commuters are included in the analysis. Wrongly administrated cases already droped out in the selection process. It is therefore assumed that all cases in the analysis here are representing real life situations. Taking that into account, the argument of selecting the 90% most probable cases to avoid falsely administrated trips is no longer a valid argument.

With that in mind, a second graphic has been drawn in Figure 6, in which all cases are included. As in the previous graphic, two dashed lines are presented to highlight the boundaries of three possible phases in which both mean travel-time ratio and mean working duration show a similar trend. Trying to distinguish those boundaries in this graph is more complex than in the one with only 90% of the most probable cases.

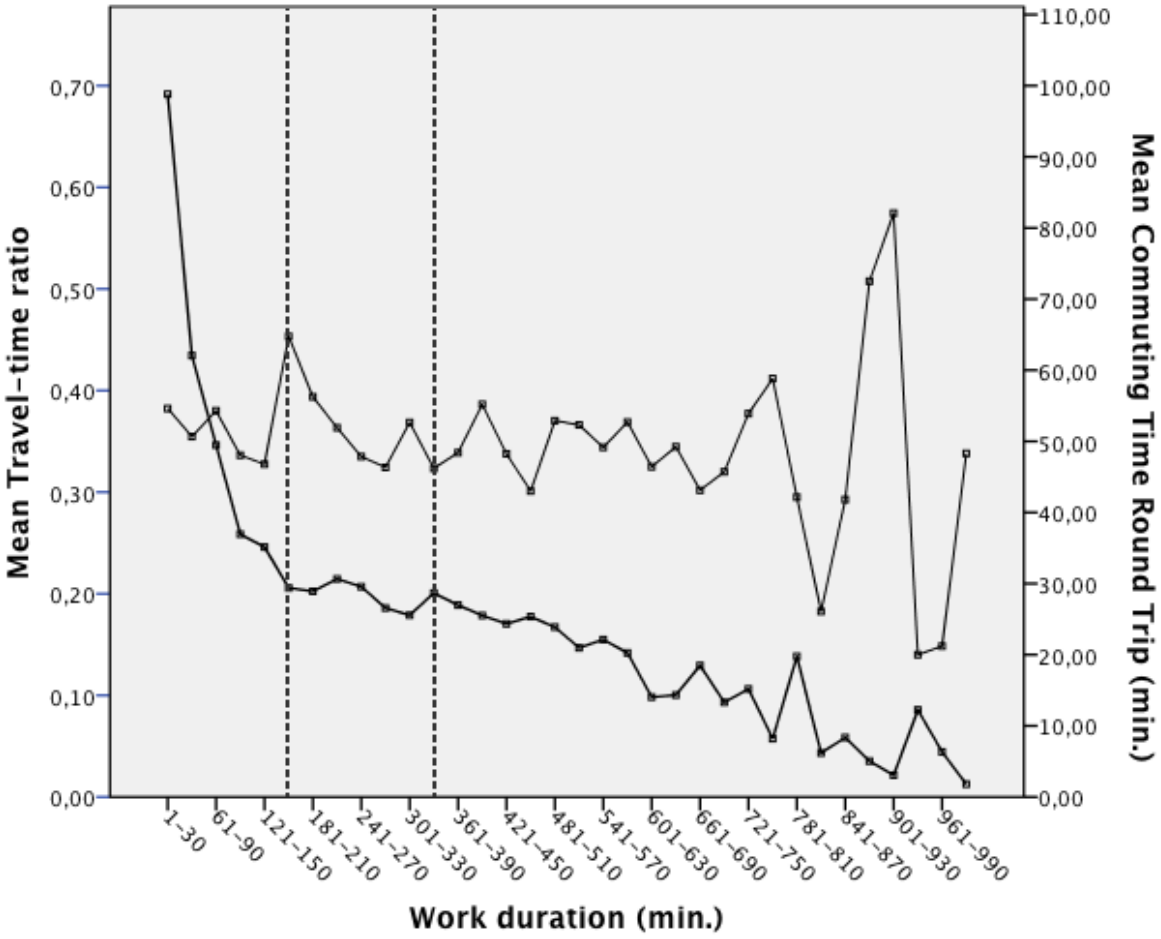

Figure 6. Travel-time ratio and work duration for all cases in 2009

What does stand out, when comparing both graphics, is that by making a limited selection it is less complex to draw conclusions than when all real life cases are included. To illustrate this, the same graphics have been drawn for 2005. In Figure 7 only the 90% most probable cases are included, Figure 8 includes all cases. Again dashed lines represent borders for different phases. The first border is at work duration of 181 minutes, as it is in all the previous graphs. But to decide where the next border should be placed in the 2005 graphics is not as straightforward as it was in the 90% selection of 2009. In 2009 the second phase consists of a relative stable travel-time ratio, due to an increasing commuting time. In 2005 however, the commuting time is not raising that clearly after the first phase. As a result of the less consistent trends in 2005, defining phases is at least not easy.

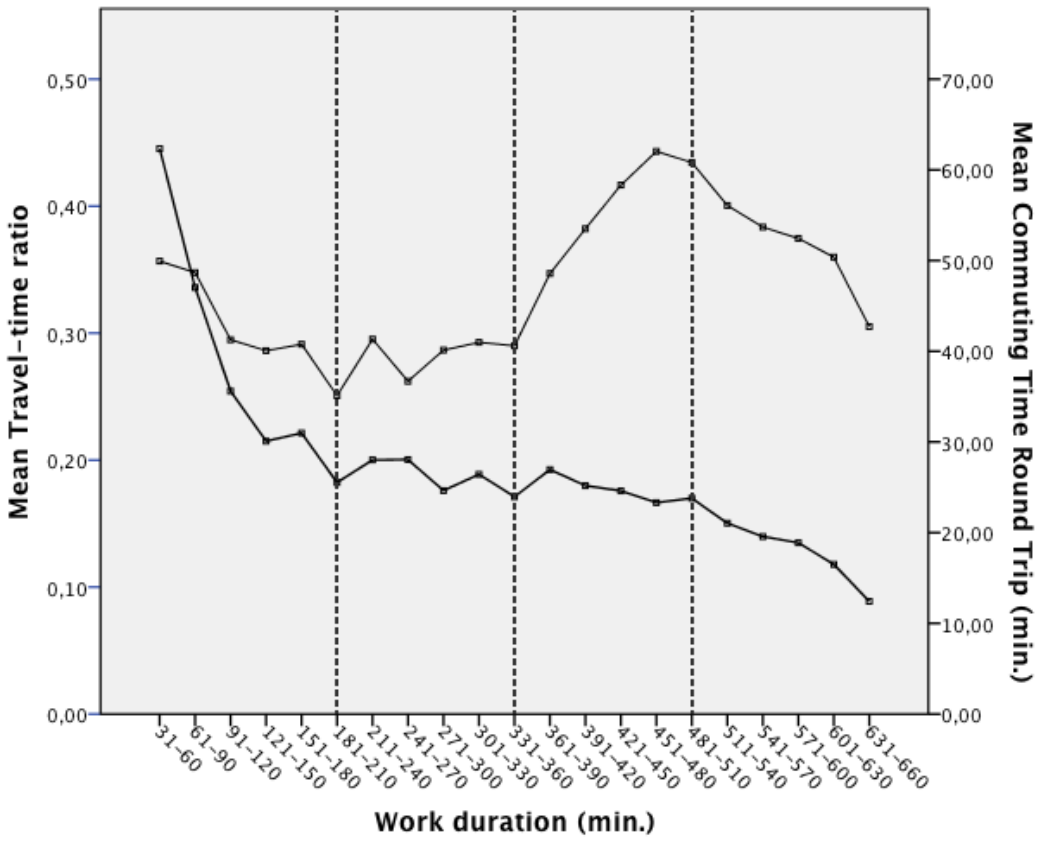

Figure 7. Travel-time ratio and work duration for 90% most probable cases in 2005

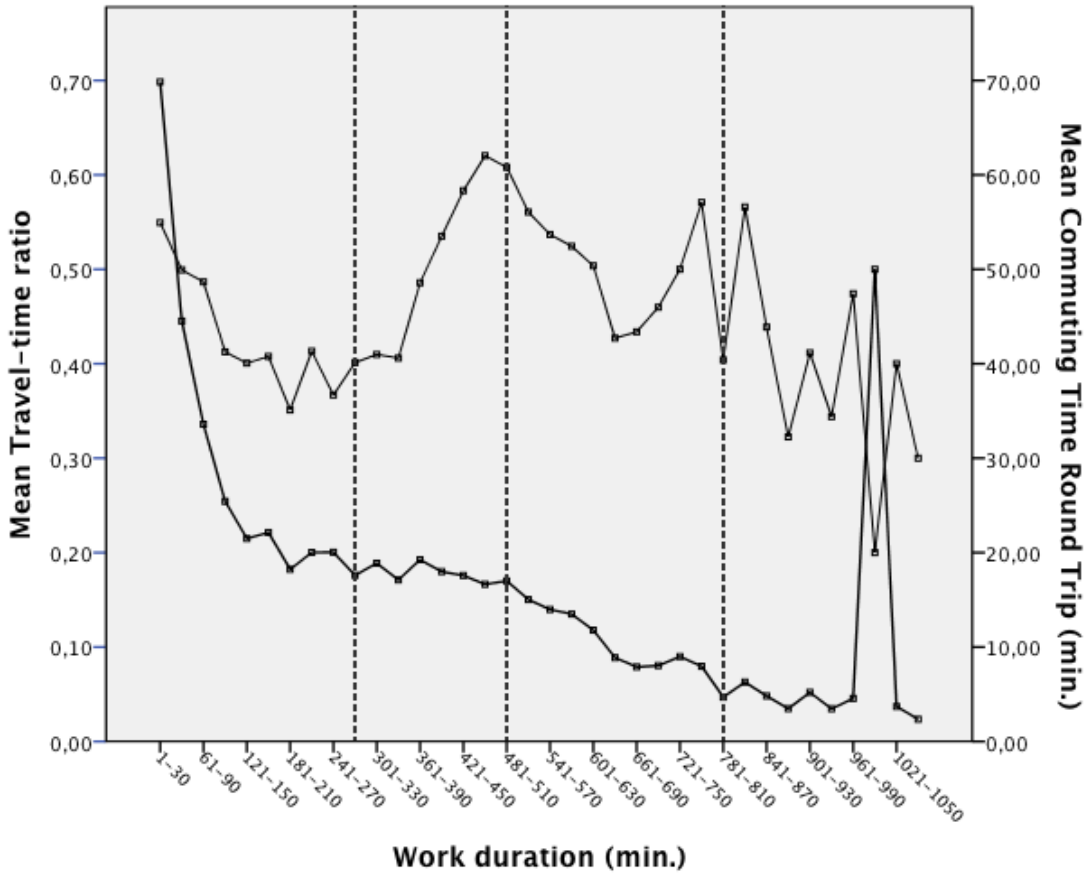

Figure 8. Travel-time ratio and work duration for alle cases in 2005

The latter stresses the role of the selections that were being made prior to the analysis. It becomes clear that the bandwidth of cases that are or are not included, has its influences on the resulting graphics and the extent to which different phases can be clearly identified. This becomes especially true when all cases are included, as has been done in this article, based on the new selection method.
The point that is being made here is not that they absolutely should or should not be included. In fact it can be argued to leave out cases that have a large proportional influence on the total, the 10% less probable cases for example. The point made here is that researchers and readers should be aware of and give notion to the implications of decisions that are being made in the data selection.
Leaving cases out might be a good decision in some travel-time researches, if augmented and noted. In other circumstance it might make sense to provide the whole picture, including every event that took place for every year and every selection possible.

Conclusions

The concept of travel-time ratios as developed by Schwanen & Dijst (2000, 2002) is a theoretical framework to help investigate the relationship between traveling times and the duration of activities. Emperical evidence for this relationship has been provided in the past, especially in an article by Schwanen & Dijst (2002) about the relationship between commuting time and work duration. That empirical evidence is given by analyses that exclude the 10% less probably data records. That in itself is a common used statistical technique to avoid so-called outliers and other false data have a effect on the analysis.

This article points out that for travel data by the Dutch National Travel Survey from 2005 untill 2009, it is possible to make a more detailed selection then a statistical selection of the 90% most probable cases. Making use of the programming languages PHP and SQL, each data record was checked to make sure that the case represents a correctly administrated trip of a classical commuter, traveling directly from home to work and vice versa.

Analysis of this data selection shows mainly four things. First of all, only half the data of commuters consists of correctly administrated trips conducted by classical commuters. Second, there are statistically significant differences in the mean travel-time ratios between 2005 and 2009. Third, from detailed graphs drawn for 2005 and 2009, it becomes clear that the trends in the travel-time ratio differ from year to year and, as a fourth conclusion, it's even harder to see real trends when using all cases selected in this paper then when only using the 90% most probable cases.

With these conclusions this paper shows that different methods of data selection can lead to different analysis. The data selection method that will be used in research can always vary, based on different needs and possibilities. But what can be noted from the result of this paper is that the influence of a data selection can influence the result of the subsequent analysis. Therefore, researchers who are investigating travel-time ratio should take notion of this effect.

This paper does not deny the relationship between working time and commuting time. Instead it has illustrated the effect of different data selections such as the year or years to analyse and the methods of case selection prior to the analysis. This illustration could be relevant for policy makers and scientist, especially those who are dealing with travel-time ratio research.

Acknowledgments

The author wishes to thank Egbert van der Zee for his enthusiasm and knowledge he shared prior and during writing of this paper. Furthermore, thankfulness is with Thijs Konijnendijk, for his constructive help during the last stages of finishing this paper. At last the department of Social Geography of the University of Utrecht has to be thanked for giving the opportunity to carry out this research.
All the opinions and views presented here are those of the author and faults can only be accounted to him.

References

Chen, Cynthia & Patricia L. Mokhtarian (2006), Tradeoffs between time allocations to maintenance activities/travel and discretionary activities/travel. Transportation 33(3), pp. 223-240.

Dijst, Martin & Valibor Vidakovic (2000), Travel time ratio: the key factor of spatial reach. Transportation 27(2), pp. 179-199.

Lyons, Glenn & John Urry (2005), Travel time use in the information age. Transportation Research Part A: Policy and Practice 39(2-3), pp. 257-276.

Mokhtarian, Patricia L & Cynthia Chen (2004), TTB or not TTB, that is the question: a review and analysis of the empirical literature on travel time (and money) budgets. Transportation Research Part A: Policy and Practice 38(9-10), pp. 643-675.

Ministerie van Verkeer en Waterstaat/Rijkswaterstaat (2009), Mobiliteitsonderzoek Nederland 2009 Het onderzoek. Den Haag.

Recker, W. W., M. G. McNally & G. S. Root (1986). A model of complex travel behaviour: Part I-Theoretical development. Transportation Research Part A: General 20 (4), pp. 307-318.

Susilo, Yusak O. & Martin Dijst (2009), How far is too far? Travel time ratios for activity participations in the Netherlands. Transportation Research Record 2134(1), pp 89-98.

Schwanen, Tim & Martin Dijst (2002), Travel-time ratios for visits to the workplace: the relationship between commuting time and work duration. Transportation Research Part A: Policy and Practice 36(7), pp. 573-592.

Wang, Donggen & Fion Yuk Ting Law (2007), Impact of Information and Communication Technologies (ICT) on time use and travel behaviour: a structural equations analysis. Transportation 34(4), pp. 513-527.

Xuedong, Lu & Eric I. Pas (1999), Socia-demographics, activity participation and travel behaviour. Transportation Research Part A: Policy and Practice 33A (1), pp. 1-18.