THE POTENTIAL OF REGIONAL PANEL SURVEYS: A COMPARATIVE ANALYSIS OF THE FRANKFURT/RHINE-MAIN PANEL AND NATIONAL HOUSEHOLD TRAVEL SURVEYS. A WORKING PAPER

Robert Schoenduwe, Martin Lanzendorf

Goethe-Universität Frankfurt/Main, Institut für Humangeographie, AG Mobilitätsforschung, Robert-Mayer-Str. 8, D-60325 Frankfurt/Main, (schoenduwe@em.uni-frankfurt.de)

DRAFT PAPER[1]

Abstract:

A continuous data collection is a key issue for the planning of new transport infrastructures and mobility options. For example transport models and recommendations for an improved transport policy depend on up to date and reliable input data. However, with the heterogeneity of the regional development scholars and transport planners increasingly emphasize the need for more diverse regional data sets (MADRE ET AL., 2006, KAGERBAUER and MANZ, 2009). The regional mobility panel in the Frankfurt/Rhine-Main area, originating from 2007, is one attempt to serve this purpose. This paper scrutinizes whether regional panel surveys are as advantageous as often stated, taking the "Rhine-Main Panel" (RMP) as its case study.

In Germany household travel data is available from three types of nationwide surveys. Firstly, the national travel survey "Mobility in Germany" (MiD, formerly KONTIV) collects structural and behavioral data regarding mobility. In order to analyze the different regions more in-depth the sample size was increased for the Federal State of Hesse. Secondly, the city focused survey "Mobility in Cities - SrV" has been conducted three times in Frankfurt/Main since 1998. Although the main survey goals differ, these two surveys have been designed to allow the comparison of key elements. Finally, in contrast to these two cross-sectional surveys with large sample sizes, the "German Mobility Panel" (MOP) only has a relatively small sample size. The MOP is a combination of a rotating panel and a one-week longitudinal survey. Panel data allows for direct observations of travel behavior changes and to analyze causal relationships related to these changes. However, the sample size of the MOP is not high enough to study spatial variations adequately.

In this paper we compare objectives, survey methods, sample methods and weighting as well as collected mobility indices of the RMP with the regional add-on of the MiD (for the years 2002 and 2008), the SrV (2008) and the MOP. Benefits of regional panel surveys as well as possibilities to combine and harmonize the data sets are pointed out. Additionally we examine to what extend the RMP data can be combined with spatial characteristics. First empirical results will be presented at the WCTR conference.

Keywords: *regional panel survey, national household travel surveys*

[1] Please do not cite without the consent of the authors

Introduction

Up to date and comprehensive mobility data is the essential basis for transport related decision making and planning. An important methodological development in recent years has been the study of behavioral change at individual level with the aid of longitudinal studies and panel data. Panels are repetitive surveys of the same person - usually annually - which aim to document behavioral changes and link these with individual and social context changes. While longitudinal and panel data offers a wide range of additional opportunities for analysis, cross-sectional surveys have the benefit of large case numbers. Moreover, due to the great research effort involved, panel surveys cannot be undertaken with sample sizes similar to e.g. cross-sectional surveys as part of national household travel surveys (i.e. Mobility in Germany 2008). Thus, the different survey methods serve complementary purposes. Although scholars as well as transportation companies (BIHN, 2002) have emphasized the benefits of regional panel surveys, only relatively few of these have been conducted so far. This is definitely linked to the greater effort entailed which is in turn often associated with higher costs. In the end, however, there is a lack of convincing arguments which justify the (alleged) increase of effort: "… the most fundamental problem has been that, as a profession, we have not been able to provide sufficient evidence on why panels are better suited to answering the policy questions of today, to convince funding agencies to fund such an undertaking" (MURAKAMI ET AL., 2006:400).

The benefits of improved mobility data for regional planning tasks have been recognized in the Frankfurt Rhine-Main region. In 2006 a pilot study examined the feasibility of a regional panel study and from 2007 onwards annual panel surveys have been conducted. In 2008 three additional mobility surveys have been conducted alongside the Rhine-Main-Panel (RMP). Therefore, our study has been able to draw on the data-sets of four household travel surveys regarding the travel behavior of the population in the Frankfurt Rhine-Main region. The 2007 RMP-wave has been employed as a reference data set. It will be examined whether it is possible to compare or supplement RMP data with the data of the surveys "Mobility in Germany" (additional sample for the federal state of Hesse) (MiD), "Mobility in Cities – SrV" (Frankfurt/Main) (SrV) and the "German Mobility Panel" (MOP) for further analyses.

This study consists of four sections. First the main characteristics of the data sets employed will be highlighted. This includes a brief portrayal of their main differences and of the benefits and disadvantages of panel and cross-sectional surveys. Subsequently, the characteristics and parameters of the surveys employed will be compared. Based on these the study will identify and explain possibilities for the harmonization and combination of the data sets employed. Finally, possibilities for the application of combined data sets will be depicted.

General remarks on the characteristics of panel- and cross-sectional surveys

The data sets available differ in terms of survey design. The MiD and SrV surveys gather cross-sectional data regarding mobility behavior on a specific day. In contrast, the MOP and RMP surveys are designed as a combination of panel and longitudinal survey. Cross-sectional surveys gather specific values of persons with particular characteristics at a certain point in time. Panel surveys, on the other hand, ask persons to provide information on certain circumstances at pre-determined intervals in time. Advantages and disadvantages of panel

surveys in comparison to cross-sectional surveys are well-known and have been described extensively (e.g. ZUMKELLER ET AL., 2006, GOODWIN, 1997, KITAMURA, 1990). The following section gives a brief summary of the most significant differences.

The cross-sectional surveys MiD and SrV, which have been conducted in intervals of several years, provide a very specific depiction of particular moments in time. They provide information on the amount of travel which took place at the time of the survey. Longitudinal surveys with a panel design such as RMP and MOP allow scholars to draw conclusions on causal links of developments in the area of travel behavior. For example, do always the same households lack access to a personal car? Or is life without a passenger-car a temporary phenomenon? Questions such as these should be at the forefront of inquiries into the reasons for limited access to transportation means. Using cross-sectional surveys for research into the dynamics of travel behavior is only suitable under specific circumstances and in a limited manner. For example, if the survey design includes retrospective questions or when data is grouped into cohorts and these are subsequently analyzed as a pseudo-panel. Cross-sectional surveys allow the identification of net changes and trends. This is true for surveys such as MiD and SrV which basically maintain the same survey methods. In contrast longitudinal data is a far better source for testing hypotheses regarding particular time-periods and for examining changes. But cross-sectional and longitudinal surveys do not exist in competition with one another; rather, they complement each other. Table 1 depicts a comparison of the advantages and disadvantages of panel and cross-sectional surveys in light of fourteen different criteria (these are not ordered according to any particular hierarchy).

Due to the great effort and commitment required of persons taking part in a panel survey, the subjects should not be chosen at random. Moreover, due to the greater organizational effort that panel surveys entail, their sample sizes are smaller than those of cross-sectional surveys. Additionally, panel surveys are subject to certain "wear-and-tear" effects like panel conditioning or panel mortality (cf. ZUMKELLER ET AL., 2006:372) because at least a few subjects tend to leave the panel over time. All of these factors diminish the **representativeness (1)** of the panel data. It is only possible to speak of accurate representation regarding of a particular population if all units have the same chance of being selected or if this chance can at least be mathematically determined (STENGER, 1994). However, these limits of representativeness in panel surveys are not relevant for the analysis of cause-and-effect relationships and direction of effects. Rather, cross-sectional and panel survey data sets are able to complement each other in order to strengthen result validity. On the one hand, it is possible to use cross-sectional data to determine to what degree the distribution of panel survey samples is an accurate representation of a certain population. On the other hand, data gathered through panel surveys can be employed to analyze the causality of changes, which cannot be analyzed by one cross-sectional dataset alone.

One of the causes for a low response-rate in cross-sectional surveys as well as panel surveys are people's attitudes towards the **protection of personal data (2)**. The continuous decrease of coverage rates highlights the gravity of this issue. Moreover, panel surveys require particular sensitivity when new subjects are recruited. Panel data cannot be gathered and stored anonymously for the duration of the period of time in which the subject takes part in the survey. Personal data may need to be collected and stored in a manner in which it can be linked to a particular subject over the course of several years. On the other hand, partak-

ing in a survey over several years may allow the interviewer and the interviewee to build a relationship of trust leading to a different kind of data disclosure.

While cross-sectional surveys require a cost-intensive recruitment of subjects for every new survey wave, panel survey **costs (3)** can be expected to decrease because the subjects of the sample remain the same. This advantage of panel surveys is not yet sufficiently recognized. Rather, an important reason why panel surveys are so uncommon in travel research is that the funding has to be secured in advance for a relatively long period of time. However, the argument that panel surveys decrease in costs over time in comparison to cross-sectional surveys pales if one takes into account the effort which is required for the analysis of panel data. The kind of survey-questions and data retrieved in panel surveys are more complex than those of cross-sectional surveys. Thus, the analysis of panel data requires more time and effort. Generally speaking, while the gathering of panel data may cost comparatively less, the analysis of panel data tends to cost comparatively more, thus, re-increasing the total overall cost of panel surveys.

The decline in the willingness to report ("wear-and-tear" effects) mentioned above negatively impacts the **reliability (4)** of the gathered data. Due to a decrease in motivation to participate people stop reporting between panel waves, which is called mortality, or stop reporting from one day to the next in multi-day surveys, which is called fatigue effect. However, generally enough context information exists, on the basis of which corrections can be made. Cross-sectional surveys have to be corrected at times as well. For instance, non-response-problems are a main source of negative impact on the reliability of cross-sectional survey data. However, the correction of such inaccuracies generally tends to be more difficult as only little context information is available. The issue of data reliability and the possibility of correction in turn affect the **validity of results (5)**. Results of aggregated analyses of panel data tend to show greater validity due to the better possibilities of making corrections and of weighting data on the basis of secondary statistics. However, due to the small sample sizes disaggregated analyses are problematic. In contrast, results of aggregated analyses of cross-sectional studies show less validity while the validity of ratio values is high. The survey characteristics mentioned so far are also relevant for the calculation of **data quality and completeness (6)** as well as for marking missing values. If one combines panel and longitudinal surveys over the course of one week in the RMP and MOP redundancies of data occur which allow a control of the quality of gathered the data. On the other hand, when surveys conducted on a single day less context information is available. This increases the necessity to make a larger amount of unfounded assumptions.

If we consider the possibilities for analysis, the advantage of a complementary use of panel- and cross-sectional data becomes evident. It is possible to calculate **net changes (7)** on the basis of a sample if several waves of panels and cross-sectional surveys have been conducted at certain temporal intervals. For this purpose net changes are defined as changes of the margin sums of two cross-sections. At this point it may be necessary to take into account existing limits of representativeness that may occur in panel surveys. It is possible to combine several panel waves to increase the sample size. In the past the calculation of net changes on the basis of cross-sectional data have often lead to the assumption of stability. However, behind this supposed stability at the collective level lye individual developments that can only be examined on the basis of panel data. Thus, the calculation of **gross**

changes (8) is able to reveal that minimal net changes may in fact express a compensation of oppositional effects. Certain fluctuations in demand can be far better explained if one is aware of gross changes. Thus, the knowledge of gross changes can help to identify target groups for measures in the area of mobility. Moreover, this kind of knowledge can generally be advantageous for modeling. This is particularly true in light of the common criticism that traffic models are general based on the assumption that processes and measures are reversible. However this kind of reversibility has been disproven, for example, in studies on the access to and choice of transportation means (GOODWIN, 1977, BLASE, 1980).

Table 1 – Advantages and disadvantages of panel- and cross-sectional surveys - *Adopted from: ZUMKELLER ET AL. (1997), HENSHER (1987),DAVIES UND PICKLES (1985)*

CHARACTERISTICS	PANEL	CROSS-SECTION
Survey design and data quality		
1. Representativeness	o	++
2. Privacy concerns	+	+
3. Dynamic of costs	+	o
4. Reliability of data	o	o
5. Validity of results	+-	+-
6. Data quality and completeness, Handling of missing values	++	o
Possibilities of analysis		
7. Calculation of net changes	+	++
8. Calculation of gross changes	++	-
9. Variability of interpersonal behavior	++	o
10. Variability of intrapersonal behavior during the course of a week	++	-
11. Likelihood to register long distance journeys	++	+
12. Conduction of before and after studies	+	o
13. Registration of seasonal differences in travel behavior	-	++
14. Registration of regional differences	(++)	++

The use of longitudinal data in studies regarding the **variability of interpersonal behavior** (9) increases the reliability of results. Cross-sectional surveys may give rise to random effects because they provide data on the mobility on the actual day on which the survey is conducted. In contrast, longitudinal data allows the attribution of "normal" behavior referring to information of the whole week. When cross-sectional data gathered on a single day is used, the total variation is usually attributed to interpersonal behavior differences. In contrast, the use of longitudinal data makes it possible to divide the total variation into an interpersonal and an intrapersonal part. Thus, it is possible to determine the **variation of a single person's behavior during the course of a week** (10). This is not possible with sample data collected on a single day. Moreover, the **likelihood to register long distance journeys** as well as trips covering several days (11) is greater when conducting panel surveys because the mobility behavior of a whole week as well as the routines of specific weekdays are

known. This kind of likelihood is generally lower when sample surveys are conducted on a single day. The attempt to compensate for this in cross-sectional surveys includes retrospective questions regarding travels as well as data collection on the normality of the day the survey takes place. Finally, panel data allows the **conduction of before and after studies** (12) with smaller sample sizes. In contrast, the possibility to make significant statements on the basis of cross-sectional data in regards to post events or post measure changes are only possible if a large number of subjects display behavioral changes.

One advantage of large-scale cross-sectional surveys is the **registration of seasonal differences in travel behavior** (13). These differences are registered because the days the survey is conducted are usually distributed over the course of a whole year. Panel-surveys are often conducted in the autumn months. At this time of year holidays and the weather are expected to have the least amount of impact on travel behavior. However, as RÜMENAPP and STEINMEYER (2005)have shown, it is difficult to justify this choice by statistical means. The **registration of regional differences** (14) would be difficult on the basis of a national panel-survey due to the small sample-size of panel surveys. However, cross-sectional surveys are able to provide sufficiently precise values for single regions. Panel surveys allow regional comparisons if regional panel surveys are conducted.

As this overview illustrates cross-sectional surveys such as MiD and SrV and panel-surveys such as MOP and RMP should not be understood as substitutes of one another but as complements to each other. On the one hand, cross-sectional survey data can be used as key survey in order to identify trends and to delineate the frameworks of developments in the transport sector. On the other hand, the MOP and RMP data can be used to gain an overview of ongoing short-term changes to determine the cause-and-effect relations of such changes. A comparison of panel data with cross sectional surveys can serve to control and steer panel samples as well as to determine factors for extrapolation and weighting (cf. RAIMOND, 2008). The benefits of the complementary usage of panel- and cross-sectional surveys have already often been emphasized. However, this method has so far only seldom been put into practice in actual research and planning (e.g. WITTWER, 2008). The kind of problems which need to be anticipated in the combined analysis of panel und cross-sectional data will be highlighted in the following section.

Comparison of methodological aspects between selected transport surveys

Besides the RMP, the MiD, SrV and MOP also provide data on the mobility of the residents in the region of Frankfurt Rhine-Main. All of these surveys cover the years 2007- 2009. As the surveys are congruent in terms of temporal frame and research topic, the question arises whether the gathered data is compatible and to what extent it can be combined. In this section we compare objectives, populations, survey methods, sample methods and weighting employed in the RMP, the regional add-on of the MiD (for the years 2002 and 2008), the SrV (2008) and the MOP. This analysis is the foundation for further evaluations in this paper. In order to arrive at a qualified comparison of different travel surveys it is necessary to closely examine how the surveys content and methods differ. Survey characteristics such as the composition of samples, the survey method and the survey design have a great influence on the quality of gathered data. Therefore, it is necessary to interpret and compare such data in the context of the method that has been employed for data collection. First we describe and

compare the individual data-sets before considering the possibility of harmonizing and combining these different kinds of data.

Survey populations and survey periods

Firstly, the travel surveys on which this paper draws will be briefly introduced and their individual goals will be delineated (c.f. Table 2). While these surveys share certain core elements they differ regarding their individual aims. The first wave of the RMP survey was conducted in autumn/winter of 2007/2008. The RMP survey methods and documents draw on the long-term experience with the MOP. A pilot study was conducted in the year 2006 and has also been included in this paper's data-set. The RMP gathers regional travel data of residents in the Frankfurt Rhine-Main region. It aims to estimate regionally specific developments in travel behavior. The MOP has been conducted annually since 1994. The MOP aims to monitor and analyze long-term developments in national German travel behavior. The MOP places a special emphasis on the understanding of individual diversity in travel behavior. This includes a particular focus on the temporal sequence and the direction of cause-and-effect-relationships associated with changes in travel behavior.

The MiD-survey 2002 follows in the foot-steps of the KONTIV[2] survey of travel behavior conducted in 1976, 1982 and 1989 in Western Germany. With the MiD 2002 a change in the survey design has been established. The follow up of cross sections was continued in 2008. Overall, the MiD aims to gather data on everyday mobility which is representative of German residents. A comprehensive catalogue of questions allows an examination of diverse spatial and thematic relationships. This paper draws on the 2002 and 2008 data-sets from the regional add-on for the federal country of Hesse. The purpose of the add-on is to allow more in-depth, sub-regional analysis of residents' mobility. However, generally speaking, the MiD supplies federal and national planning-levels with mobility data regarding regional cross-sectional values, seasonal variations, and values specific to area types. In contrast, the SrV largely provides data regarding communal issues. The SrV aims to provide specific values regarding average work day travel in individual cities and municipalities. Thus, the survey is able to provide a database for local travel policy and planning. The 2008 SrV data for Frankfurt/Main is included in this paper.

Despite the differences mentioned so far, all the surveys share the same fundamental goal: they seek to attain knowledge on whether people are mobile within certain time periods, what kind of mobility resources people have access to, and in what manner this mobility is realized. The survey population, meaning the number of people about whom such statements are to be made, is determined by spatial, temporal and thematic boundaries.

Although the basic population of the surveys employed here are similar, they also differ to some extent. Firstly, the surveys differ in terms of spatial delineation. The RMP survey area encompasses the Frankfurt Rhine-Main region including the county-free cities of Darmstadt, Frankfurt, Mainz, Offenbach, Wiesbaden as well as municipalities from the counties Darmstadt-Dieburg, Groß-Gerau, Hochtaunuskreis, Mainz-Kinzig-Kreis, Main-Taunus-Kreis, Offenbach and Wetteraukreis. This paper chooses the RMP survey area as the area of reference. The other surveys will be adapted as much as possible. The MOP survey encom-

[2] "Kontinuierlichen Erhebungen zum Verkehrsverhalten" (ongoing survey of travel behaviour)

passes West-German households since 1994 and is conducted nation-wide since 1999. The MiD survey is conducted nation-wide as well. The regional add-on for the federal state of Hesse encompasses Hessian municipalities and the city of Mainz (MiD 2008 only). The 2008 SrV survey has been conducted in 76 cities and municipalities. This paper draws on the SrV's Frankfurt/Main sample. All survey subjects are residents of the delineated survey areas. Thus, transit travel has been totally excluded. The basic population has additional limitations which partially result from the methods of sample selection (see below).

Table 2 – Characteristics of basic populations and survey time frames

CHARACTERISTIC	RMP	MOP	MID (2002, 2008)	SRV 2008
Basic population	Residential population of the Frankfurt Rhine-Main region („Planungsraum Frankfurt Rhein-Main)	Residential population of Germany with German-speaking householder (until 1999 persons of West German federal states only) (households with landline only)	Residential population of the federal state of Hesse	Residential population of Frankfurt/Main
Reference for the basic population	Phone book (only persons with address information specified)	Access-panels (generated in telephone interviews covering multiple topics)	Municipal census "Einwohnermelderegister" (EMR)	Municipal census "Einwohnermelderegister" (EMR)
Age of respondents	Persons from the age of 11 on	Persons from the age of 11 on	All persons (Persons older than the age of 14 were used for sampling)	All persons
Field time	12.09.2007 - 28.03.2008 (first wave)	Since 1994, each autumn	07.12.2001 - 22.12.2002 January 2008 - April 2009	01.01.2008 – 31.12.2008
Number of days surveyed per person (Distribution over a week)	Mo. – So. (every respondent reports one week; with one seventh starting each week day)	Mo. – So. (every respondent reports one week; with one seventh starting each week day)	Every respondent reports one day out of all possible days of survey	Respondents report Tuesday to Thursday ("middle workday"), holidays excluded
Definition of a complete household for the net sample	Test of plausibility by using a case-to-case approach	Test of plausibility by using a case-to-case approach	Interviews of at least 50% of the household members available	Interviews of all household members available

All the surveys use household samples. RMP and MOP require all household members from the age of 11 to write a daily travel diary. The RMP does not explicitly limit in terms of subject nationality. However, the survey documents are written in German which at least requires subjects to have a working knowledge of the German language. While in theory all household members should be part of the MOP survey, in practice two limitations have been introduced. Firstly, analogue to the RMP only persons from the age of 11 are able to participate. Secondly, no more than five household members (generally, the eldest five) are asked to record their mobility in a daily travel diary. In contrast, the MiD survey includes all household members in principle. Persons from the age of 14 are asked to give personal statements on their mobility on the day the survey takes place. Moreover, these household members are able to give statements in the place of younger children as well as for those who are unable to participate in person at the time of the telephone survey (proxy). However, in these proxy interviews the questionnaire is not surveyed completely. This has resulted in a high non-response rate for these characteristics. In the MiD surveys a household interview is termed

successful if it provides the complete information of 50% of the household members. In contrast, panel surveys employ a range of different criteria to examine each case study separately regarding its suitability for the sample finally used for an analysis. Following again another method the SrV survey applies a 100% rule: households are only selected for an analysis if all household members have participated in the survey.

Secondly, panel and cross-sectional surveys differ in terms of their temporal framework. Both MiD and SrV have been designed to provide information representative of travel participation throughout the year. Thus, they have been designed as year-round surveys, meaning the days of survey are distributed evenly throughout the entire survey year. Panel surveys record mobility during the autumn months. The MOP and RMP require subjects to document data regarding their every day mobility in a travel diary for the duration of seven days. The SrV aims to gather information on everyday travel behavior on work days. The SrV thus is conducted on the so-called “middle workdays” Tuesday, Wednesday and Thursday. If one of these days proceeds, follows or is itself a holiday it is exempted from the survey.

If one seeks to compare parameters of the surveys discussed here it is necessary to take into account the different survey time frames. RÜMENAPP and STEINMEYER (2005) have analyzed the impact of different survey time frames on the results of data collections for the 2002 MiD and 2003 SrV. They make recommendations for handling differences in the course of the week and the course of a year and propose certain corrective factors. Whether such corrective factors also need to be employed in the use of a combined data-set should be examined in a separate study. A comparison of central parameters so far has given the general indication that the use of such corrective factors is not necessary.

Sampling, survey design and survey methods

Criteria for the selection of samples have been chosen so that all characteristics of interest in a given sample are equally distributed to the characteristics in the basic population. This is accomplished in panel surveys through the specific selection of subjects. Households are apportioned by household size, household type (referring to the socio-economic status) and a spatial factor. This method has been chosen based on the high motivation that subjects need to possess for participating in panel surveys. After all, subjects need to participate in the panel for over three years. All panel surveys discussed here are designed as rotational panels. This means subjects remain part of the survey sample for three survey waves. Then new subjects refresh the sample and thus maintain its representativeness as best as possible.

The RMP subject selection follows a two stage process. Firstly, municipalities in the survey area are divided according to spatial categories in regard to population density and central place function (“BBSR-Raumtypen”). Secondly, subjects for the sample are drawn on the basis of phonebook entries. Thus, the basic population encompasses persons in the Frankfurt Rhine-Main region whose address is listed in the public phone book.

Similar to the RMP, the MOP recruitment process takes place in two stages. Firstly, the survey area is divided into spatial categories and municipalities are selected at random. The spatial categorization basically follows the BIK-categories (BIK 2001). Thus, the categorization slightly differs from that in the RMP. Secondly, households from the randomly chosen

communities are selected via access-panels. These households are apportioned according to household type and car ownership. The access-panels are composed in telephone interviews covering multiple topics. Persons are contacted for these surveys by Random-Digit-Dialing.

MiD and SrV draw a sample on the basis of the municipal census ("Einwohnermelderegister"). Thus, in contrast to the panel surveys the MiD and SrV truly sample at random. However, also the MiD survey sampling has several stages. Firstly, communities are selected according to spatial categories (BIK-categories, population sizes and BBR area types). These categories also differ from those in the RMP, however, they are basically comparable as the RMP-Region is largely homogenous in terms of population density and access to core cities. In the regional add-on for the federal state of Hesse the survey area was additionally divided according to county-free cities and sub-regions. Secondly, samples of individuals were requested from the municipal census of the chosen municipalities. From these household samples (household address) were determined. Although the SrV likewise samples via municipal census the SrV employs a simple random sampling method. The basic population encompasses all persons registered as residents in Frankfurt/Main.

The possible choice of survey methods includes face-to-face interviews, postal surveys, and telephone surveys. Each survey method has advantages and disadvantages which have caused controversies among scholars (cf. BONNEL and LE NIR, 1998). The survey methods employed by the surveys discussed here guarantee a high precision in registration of the requested attributes. All surveys have been preceded by comprehensive pilot studies or pretests. Additionally, researchers are able to draw on long-standing experiences. Differences in the design of the questionnaires should only have little impact on the quality of the gathered data as opposed to other sources of error like non-response issues. However, there is no scientific study which examines this thesis regarding the questionnaires employed. The data collection process of all surveys is generally divided into three phases. First the households are contacted by letter or telephone. Then the survey documents are sent to each household. Finally, in the case of non-response a reminder is given by letter or telephone. The measures taken in these stages differ in detail from one survey to another; however, we will not consider these details here.

In the RMP the sample households are at first informed via mail about the implementation of the survey. Subsequently an introductory interview serves to query the subjects' willingness to participate and to gather preliminary data on the household. During the main survey phase the participating households receive a household questionnaire as well as travel journals for every household member from the age of 11. Moreover, subjects receive an incentive for their participation. The MOP follows similar procedures. The recruitment via phone is followed by contact via post.

Table 3 - Sampling, survey design and survey methods of t he data sets used for the study

CHARACTERISTICS	RMP	MOP	MID 2002, 2008	SRV 2008
Principle	Quota-sample	Quota-sample	Two-stage stratified random sample	random sample
Selection stages	*Stage 1*: Municipalities (stratified random sample; stratified according to six BBSR[1] area types) *Stage 2:* Households, Persons above the age of 10 (phone book)	*Stage 1*: Municipalities (stratified random sample; stratified according to three to five BIK[2] area types) *Stage 2*: Households, Persons over the age of 10 (access panel); since 2001 quota-sample by household types (household size, employment status, number of children per household) and car ownership	*Stage 1*: Municipalities (stratified random sample; stratified according to BBSR[1] area types and number of inhabitants per municipality) *Stage 2*: Sample points selected at random from municipal census; home addresses selected at random from sample points	*Stage 0:* participating cities, municipalities *Stage 1*: Households (municipal census)
Additional limitations of sampling	None	A maximum of five persons per household receive a trip diary	none	none
Survey design	Postal survey (Main part); (CATI)	Postal survey	Telephone survey (CATI) + postal surey + (since 2008) online survey	Telephone survey (CATI) + postal surey + (since 2008) online survey
Incentives	Yes	No	No	No
Additional non-response survey?	No	Yes	Yes	Yes
Maximum of trips	All trips have to be recorded; as required new trip diaries can be requested by telephone	All trips have to be recorded; as required new trip diaries can be requested by telephone	Up to 8 trips per trip diary; in MiD 2008 up to 12 trips have to be recorded (Data transmission CATI only)	All trips have to be recorded; up to 8 trips per trip diary further trips separately or on a new trip diary

[1] Federal Institute for Research on Building, Urban Affairs and Spatial Development ("Bundesinstitut für Bau-, Stadt- und Raumforschung")
[2] German classification of metropolitan areas (BIK 2001)

In the MiD the selection of the Municipal census is followed by researching the phone numbers of the respective households. The households who did not have a phone number listed amounted to 40% and differed systematically from the other subjects (a higher rate of non-nationals, a higher rate of recently relocated households). Therefore, the MiD employed not only telephone surveys (CATI) but also postal surveys (PAPI). Both survey methods start off with an informational letter. Subsequently, the CATI-part is contacted via phone. If the subjects agree to participate in the survey, household data is collected via the telephone. Moreover, each household member receives a personalized travel diary which states the date of survey. The MiD 2008 included the option for subjects to enter household data online. The households were called again between one and 14 days after the date of survey and all household members were asked to pass on the collected trip data. The MiD survey allowed proxy interviews in previously defined exceptional cases. Households which provided their phone number in the first stage of the PAPI sampling (household interview) transferred to telephone interviews for the second phase (person- and trip interviews). From 2008 the MiD

survey only permits person- and trip interviews via telephone. Subjects who did not enter their phone number in the household questionnaire were not included in further interviews. While the methodological mix of postal, telephone and online survey is practical as it allows researchers to contact a large number of subjects this mix also creates particular non-response effects which we depict later.

Previous to the 2003 SrV survey and the 2002 MiD survey the survey institutes and clients agreed to attune the travel surveys as much as possible to one another. This resulted in the strong similarities between the SrV and the MiD survey process. The SrV survey is also conducted through a combination of postal, telephone survey, supplemented by an online survey.

To examine whether there is a systematic difference between participants and non-participants all surveys except the RMP conducted an additional non-response survey. For example, for the MiD this additional survey revealed that elderly people, non-nationals, and members of large households have a lower rate of participation. Moreover, a mobility bias was detected. People, for whom mobility is a highly relevant topic, e.g. frequent travelers, regular public transport customers, and members of households with two cars or more, are more likely to participate in the survey (DIW and INFAS 2003).

Sample sizes (Net-samples)

The RMP surveyed persons from 123 municipalities in twelve counties or county-free cities. Morover, the RMP is the only survey which posesses the municipality key which allows the definite identification of the participating communities. This makes the RMP the only survey which allows data analysis at the municipality-level. With the aid of the key, connections can be drawn with spatial data from other surveys and statistics of federal statistical offices. This creates the possibility for various additional kinds of analyses. In contrast, the other surveys only provide data for the next higher level, the county, for reasons of privacy concerns. The MiD does not even provide this kind of data; however, the county name can be partially reconstructed from the data the MiD collects. In the SrV all subjects reside in the city of Frankfurt/ Main.

The RMP data-set employed in this paper also includes the households from the second wave of the pilot study. In the following sentences the sample size of this wave is stated in brackets. In total 3.920 households (120 in the pilot panel) have been contacted via phone and participated in an introductory interview. Of these households 2.089 (3) refused to participate in the survey. A total of 775 (0) were reported as quality neutral non-responses as it was impossible to reach these households through the listed telephone numbers. This brought the adjusted gross sample down to a number of 1.038 (117) sample units. Of these only 540 (106) households sent back the completed survey documents which amounts to a response rate of 52%. After the identification and marking of unusable questionnaires, the net sample contained a total of 520 (106) households.

The MOP interviewed approx. 750 households per survey wave. This equates each time to approx. 1.500 persons filling out travel journals. Approximately two thirds of these subjects have participated in previous one or two survey waves. No exact numbers have been provided for the waves' gross sample size, the refusals to participate, quality neutral non-

responses and the adjusted gross sample. However, due to the particular recruitment process or, respectively, the quota-sampling of the survey participants these numbers are only of little concern. The numbers much more relevant to determine the panel's quality are repetition rates and panel mortality rates between waves. Repetition rates lye at an average of 76% (Zumkeller et al., 2008b). Table 4 depicts the values for the 2007 survey wave and the respective number of units in the survey area. The table also depicts the respective values of the other surveys. The columns MiD (2002, 2008) contains the numbers of the 2002 and 2008 MiD add-on for Hesse as well as the data of the Rhine-Main area. The numbers of recorded trips also contain trips stated with implausible values. These are excluded from further analysis.

Table 4 – Sample size (Number of units in the net sample) of the data sets used in the study

CHARACTERISTICS	RMP	MOP 2007	MID 2002	MID 2008	SRV 2008
Municipalities	123	326 k.A. (RM)	85	k.A.	1
Counties / County-free cities in the RMP-region	11	11	11	11	1
Response rate	52% (pilot study: 91%)	k.A. (repetition rate ca. 75%)	ca. 40%	ca. 20%	ca. 30%
Households	520 (pilot study: 107)	904 44 (RM)	8.761 3.293 (RM)	9.049 5.491 (RM)	2.726
Persons	1.013 (pilot study: 194)	1.567 69 (RM)	21.229 7.531 (RM)	20.902 12.078 (RM)	5.067
Trips	22.585 (pilot study: 4.679)	37.520 1.739 (RM)	57.459 20.557 (RM)	64.576 37.579 (RM)	14.266
Persons over the age of 10	1.013 (pilot study: 194)	1.567 69 (RM)	18.993 6.755 (RM)	19.471 11.281 (RM)	4.549
Trips of persons over the age of 10	22.585 (pilot study: 4.679)	37.520 1.739 (RM)	52.386 18.800 (RM)	60.359 35.188 (RM)	12.691

RM = Elements in the Frankfurt Rhine-Main Region (according to the study area of the RMP)

All of the surveys provide various weighting factors for the individual, household and trip level in order to counterbalance methodological and non-response-related biases. The weighting factors provided could not be used in the application of combined data-sets because the basic populations have been attuned to the RMP survey area. Whether the application of weighting factors is reasonable has to be decided case by case according to research question.

Harmonization and combination of data sets

In the previous sections the travel surveys in this paper have been compared in terms of survey design. If the aim is to apply the surveys' data sets comparatively and in combination with one another, one needs to take into account the differences stated above. Moreover, such an application also requires consistent definitions and references regarding surveyed

variables. For this purpose, all variables and characteristics have to be listed and categories regarding, for example, the choice of transportation means, the possession of a driver's license and the definition of trip purposes and trip characteristics have to be harmonized. Additionally, the persons to be examined need to be demarcated in a unified manner. Differences in the basic population have already been highlighted above. Additionally, for example, the definition of mobile and immobile persons and the definition of long-distance journeys are important for the calculation of mobility measures.

A range of studies already exists regarding the comparability of the data sets employed in this paper. The comparability of the MOP and RMP data was a key requirement in the RMP design. This was to create the possibility of increasing the RMP sample with data from MOP participants. For this purpose the data-sets can be combined in a number of ways. Firstly, the data sets can be combined according to spatial parameters. This may only entail considering persons who are residents in the RMP survey area or it may entail choosing subjects according to certain area types. The combination according to area types derives from the assumption that persons which reside in the same area type display similar behavior patterns. Secondly, the data sets can be combined according to temporal criteria. This method is based on the assumption that travel behavior does not undergo fundamental change in the course of a few years. Thus, several survey waves are combined to expand the sample size.

Various studies also exist regarding the comparability of the MiD and MOP as well as regarding possible biases in the MOP (Zumkeller et al., 2002). These studies have identified important differences which result from survey design. The studies generally concluded that the research methods employed by the surveys have their particular possibilities and constraints. Nevertheless, a combined use of the surveys is possible. Moreover, this combined use is able to draw on the respective advantages of each survey and, thus, is able to create a positive complementary effect. Although the MOP samples are relatively small, the regular survey pattern allows to draw a detailed image of travel demand. Zumkeller et al. (2008a) propose a concept of pooled analysis based on a combined sample. The concept makes the assumption that indicators of travel behavior does not change fundamentally over the course of three years. This assumption is based on the relatively stable socio-demographic and economic conditions existing within the sample. The combination of several years in one analysis also creates the possibility for a more differentiated spatial analysis. Results of such an analysis should be interpreted as cross-sectional means over the course of several years. The RMP is only able to provide this kind of advantage after several survey waves have been conducted. Comparisons of SrV and MiD available so far confirm that the parameters of both surveys can generally be combined for planning and modeling tasks (Wittwer 2008). However, it was also demonstrated that this would require considerable and very time-consuming extra working steps.

In the present paper the harmonization of data-sets was conducted in three stages. First the variables which exist in all surveys and, thus, are suitable for combined data-sets were identified. The categories of the variables were adapted accordingly. The variables where examined as to whether they were immediately comparable or whether they needed to be recoded. Subsequently, the data-sets were combined. The RMP data sets were used for reference. Four data-sets were created: a household data-set, a person data-set, a trip data-set, and a car data-set. The procedures employed were documented in detail (SCHOENDUWE and

LANZENDORF 2010). In the following section's several differences will be highlighted which became evident in the course of the data harmonization and which should be considered in the data analysis and interpretation.

Spatial structure and socio-demographic characteristics

In this section the samples' spatial and socio-demographic characteristics will be compared on the basis of several descriptive examples. The area type according to BBSR can be used for intraregional comparisons. This classification is a basic formation of groups most significantly on the basis of population density and the size and/or the central place function of core cities. Distinctions are made according to core cities and other counties or county-free cities together with their surrounding area ("Kreisregionen"). County-free cities with more than 100.000 residents are considered core cities. The Kreistypzuordnung of the survey area partially exhibits substantial differences. This results from the distribution of respondents according to counties (Table 5). If one compares the respondents' spatial distribution according to area types a distributional bias in favor of the county-free cities Darmstadt (MiD 2008) and Wiesbaden (MiD 2002) becomes evident. This bias is evident in the higher rate of subjects from core cities in agglomeration areas. Moreover, the 2002 MiD is missing subjects from Mainz.

Table 5 – Respondents in area types according to BBSR[1]

AREA TYPE ("KREISTYP")	**RHINE-MAIN-REGION**[1]	**RMP**	**MOP 05-07 RM**	**MID 2002**	**MID 2002 RM**	**MID 2008**	**MID 2008 RM**	**SRV 2008**
N	3.326.366	1.207	197	21.229	7.531	20.902	12.078	5.076
Core cities in agglomeration areas ("Kernstädte in Agglomerationsräumen")	35,9% (1.195.306)	23,0% (278)	28,9% (57)	14,5% (3.079)	40,9% (3.079)	28,7% (5.998)	49,7% (5998)	100% (5.076)
Counties of very high-density in agglomeration areas ("Hochverdichtete Kreise in Agglomerationsräumen")	37,8% (1.258.579)	36,8% (444)	28,9% (57)	18,2% (3.868)	31,8% (2.393)	24,8% (5.191)	36,4% (4392)	-
High-density counties in agglomeration areas ("Verdichtete Kreise in Agglomerationsräumen")	20,6% (686.420)	35,5% (428)	26,4% (52)	24,8% (5.267)	27,3% (2.059)	17,5% (3.651)	10,9% (1313)	-
Core cities in urban areas ("Kernstädte in verstädterten Räumen")	5,6% (186.061)	4,7% (57)	15,7% (31)	1,0% (213)	-	3,2% (664)	3,1% (375)	-
High-density counties in urban areas ("Verdichtete Kreise in verstädterten Räumen")	-	-		21,9% (4.645)	-	13,5% (2.818)	-	-
Rural counties in urban areas ("Ländliche Kreise in verstädterten Räumen")	-	-		10,7% (2.264)	-	8,3% (1.745)	-	-
High-density rural counties in peripheral areas ("Ländliche Kreise höherer Dichte in ländlichen Räumen")	-	-		8,9% (1.893)	-	4,0% (835)	-	-

[1] Federal Institute for Research on Building, Urban Affairs and Spatial Development ("Bundesinstitut für Bau-, Stadt- und Raumforschung")
[2] Rhine-Main-Region = "Planungsraum Rhein Main"
MOP05-07RM = pooled data set (MOP-waves 2005-2007 (municipalities in the RMP-study area only)
…RM = respondents with residence in the RMP-study area only

Table 6 depicts the descriptive comparison of the characteristics "gender", "nationality", "age" and "household size". In regard to Gender the surveys only differ slightly. In contrast, the characteristic of nationality reveals great differences between the basic population and the samples. In fact, subjects' nationality was only surveyed in the RMP and MiD 2002. In both of

these data-sets persons without German nationality are clearly underrepresented. Additionally, the MiD data-set is problematic because of a relatively big amount of proxy-interviews. This problem will be examined more closely in the following section. Concerning the respondents age the 2002 MiD's low rate of persons above 65 is especially evident. Moreover, the RMP's and MOP's limitation to persons from the age of 11 becomes explicit. Finally, in all of the surveys single-person-households are underrepresented while two-person-households are overrepresented in reference to the basic population.

Table 6 – Sociodemographic characteristics of respondents

CHARACTERISTICS		RHINE-MAIN-REGION[1]	RMP	MOP 05-07 RM	MID 2002	MID 2002 RM	MID 2008	MID 2008 RM	SRV 2008
	N	3.326.366	1.207	197	21.229	7.531	20.902	12.078	5.076
GENDER									
Male		48,9% (1.627.150)	47,2% (570)	49,2% (97)	49,1% (10.434)	48,5% (3.655)	49,3% (10.306)	49,2% (5.942)	49,2% (2.495)
Female		51,1% (1.699.216)	52,0% (628)	50,8% (100)	50,8% (10.783)	51,4% (3.873)	50,7% (10.596)	50,8% (6.136)	50,8% (2.581)
NATIONALITY									
German		84,9% (2.827.411)	97,4% (1.176)	-	93,3% (5301)	96,0% (15.007)	-	-	-
Other nationality		15,1% (498.955)	2,6% (31)	-	6,7% (382)	4,0% (633)	-	-	-
Missing values		-	-	-	24,5%* (1.848)	26,3%* (5.589)	-	-	-
AGE (Years)									
Up to 5		5,5% (183.653)	-	-	6,0% (1.269)	6,4% (481)	3,5% (741)	3,4% (411)	6,1% (311)
6 – 17		11,5% (382.623)	9,5% (115)	5,1% (10)	15,5% (3.297)	14,2% (1.069)	13,8% (2.892)	12,8% (1.551)	13,2% (670)
18 – 29		14,1% (469.187)	10,1% (122)	8,6% (17)	11,6% (2.470)	11,1% (834)	10,9% (2.284)	10,6% (1.281)	10,5% (535)
30 – 49		31,6% (1.050.308)	34,9% (421)	44,2% (87)	32,2% (6.836)	32,0% (2.408)	26,3% (5.503)	25,8% (3.120)	30,8% (1.564)
50 – 64		18,6% (619.159)	22,6% (273)	19,8% (39)	21,6% (4.583)	22,7% (1.711)	24,8% (5.183)	25,0% (3.018)	21,5% (1.089)
65 and older		18,7% (622.166)	21,4% (258)	22,3% (44)	13,0% (2.755)	13,6% (1.024)	20,3% (4.250)	22,0% (2.663)	17,8% (901)
Not applicable		-	1,5% (18)	-	0,1% (19)	0,1% (4)	0,2% (49)	0,3% (34)	0,1% (6)
HOUSEHOLD SIZE	**N**	2.849.000[2]	626	**123**	8.761	3.293	9.048	5.491	2.249
Single-person households		35,4% (1.008.000)	26,8% (168)	35,8% (44)	15,0% (1.313)	17,9% (591)	18,9% (1.714)	22,2% (1.221)	30,9% (696)
Two-person households		34,8% (991.000)	39,0% (244)	37,4% (46)	39,6% (3.465)	42,0% (1.382)	43,7% (3.954)	44,8% (2.457)	36,1% (811)
Three-person households		14,4% (409.000)	13,7% (86)	14,6% (18)	18,1% (1.583)	16,8% (553)	16,3% (1.479)	15,0% (824)	14,5% (327)
Household with four or more persons		15,4% (440.000)	20,4% (128)	12,2% (15)	27,4% (2.400)	23,3% (767)	21,0% (1.901)	18,0% (989)	18,5% (415)

[1] Rhine-Main-Region = "Planungsraum Rhein-Main"
[2] Statistical state office of the federal state of Hesse 2006 – Figures for the Federal state of Hesse (As of 31.12.2004)
MOP05-07RM = pooled data set (MOP-waves 2005-2007 (municipalities in the RMP-study area only)
…RM = respondents with residence in the RMP-study area only

Table 7 – Descriptive statistics on respondents education and employment

CHARACTERISTICS	RMP	MOP 05-07 RM	MID 2002	MID 2002 RM	MID 2008	MID 2008 RM	SRV 2008
N	1.207	197	18.960	6.813	15.830	9.297	5.076
EDUCATION							
Secondary general school ("Volks-/Hauptschule")	16,8% (203)	14,2% (28)	27,1% (5142)	23,8% (1.620)	20,2% (3200)	17,1% (1.591)	16,3% (828)
Intermediate school ("Realschule, Mittlere Reife")	28,1% (339)	31,5% (62)	25,4% (4822)	24,7% (1.683)	25,9% (4103)	24,4% (2.269)	20,7% (1.049)
Upper secondary school ("Abitur") **and Higher education** ("Fachhochschule, Universität")	42,4% (512)	45,7% (90)	23,7% (4404)	27,9% (1.902)	33,3% (5266)	39,2% (3.645)	41,4% (2.100)
Other	-	-	0,6% (119)	0,7% (48)	0,6% (97)	0,7% (63)	-
Did not finish school (yet)	10,4% (126)	4,1% (8)	23,1% (4379)	22,4% (1.528)	19,7% (3125)	18,4% (1.710)	20,5% (1.041)
Not applicable	2,2% (27)	4,6% (9)	0,5% (94)	0,5% (32)	0,2% (39)	0,2% (19)	1,1% (58)
Missing values	-	-	10,7% (2.269)	9,5% (718)	24,3% (5.072)	23,0% (2.781)	-
EMPLOYMENT							
Full-time employed	36,2% (437)	40,6% (80)	27,1% (5.743)	28,1% (2.115)	29,8% (6.221)	30,4% (3.666)	32,9% (1.670)
Part-time employed	14,3% (173)	13,2% (26)	10,6% (2.249)	10,2% (770)	13,0% (2.719)	12,1% (1.466)	10,8% (550)
Not employed	1,9% (23)	4,1% (8)	1,9% (403)	2,0% (153)	1,4% (298)	1,5% (185)	1,9% (97)
Not employed: Housewife/Househusband	6,7% (81)	5,6% (11)	8,2% (1.740)	7,8% (585)	8,1% (1.688)	7,7% (930)	5,9% (301)
Retirement	24,4% (295)	25,4% (50)	15,5% (3.298)	16,9% (1.276)	21,6% (4.512)	23,1% (2.789)	21,8% (1.108)
Other	-	-	0,9% (195)	1,0% (76)	2,4% (507)	2,5% (302)	0,1% (7)
Not applicable	1,3% (16)	2,5% (5)	0,2% (32)	0,1% (9)	0,1% (12)	-	0,4% (19)
Student (general school)	11,0% (133)	8,6% (17)	14,7% (3.127)	13,9% (1.047)	15,0% (3.133)	14,1% (1.705)	14,4% (729)
In occupational / professional education	1,7% (21)	-	1,9% (394)	1,5% (114)	2,2% (462)	1,9% (233)	1,8% (90)
Student (University)	2,3% (28)	-	1,8% (373)	1,9% (146)	2,7% (567)	3,0% (365)	3,4% (174)
Infant in kindergarten	-	-	3,5% (738)	3,6% (271)	2,4% (504)	2,4% (294)	6,5% (331)
Infant not in kindergarten	-	-	3,0% (641)	3,2% (244)	1,3% (279)	1,2% (142)	-
Missing values	-	-	9,6% (725)	10,8% (2.296)	-	-	-

[1] Rhine-Main-Region = "Planungsraum Rhein Main"
[2] Statistical state office of the federal state of Hesse 2006 – Figures for the Federal state of Hesse (As of 31.12.2004)
MOP05-07RM = pooled data set (MOP-waves 2005-2007 (municipalities in the RMP-study area only)
…RM = respondents with residence in the RMP-study area only

Table 7 depicts the comparison of the characteristics "education" and "employment". If we consider "education" the high number of non-responses in the MiD survey is evident again. Moreover, the survey also has a noticeably low rate of participants with secondary or tertiary education. In the panel surveys the exclusion of people below the age of eleven impacts the characteristic of education. Apart from these differences, the subjects' educational profile is

similar in all surveys. Regarding the subjects' employment especially the low rate of unemployed people stands out.

Item-nonresponse

The quality of completed surveys is usually evaluated by the amount of missing values in regard to particular characteristics. The response rate is not automatically linked with data quality. If some of the respondents do not complete the questionnaires or even leave out single questions this may substantially distort the net sample in comparison to the basic population. Therefore, the extent to which results can be generalized varies according to research question and/or the employed variables. The MiD-survey poses the problem that questions in CATI and PAPI surveys have been differentiated differently. Moreover, several sets of questions were not surveyed in proxy interviews (comp. table 9). For example, questions regarding car use and the ownership of public transport season tickets are not included in the child questionnaires. These kinds of limitations also need to be made for panel surveys in the case of analysis. If one views the proxy-data in the MiD-survey as item-nonresponse the panel surveys possess a throughout better response quality. The SrV-survey surpasses MOPs and RMPs response quality due to consistently low defaults. Only variables regarding homogenous groups have a higher quality in the MiD compared to the other surveys. Questions concerning mobility limitation have only been included in the MOP since 2004. This explains the large amount of missing values. Moreover, several years of the MOP show a high nonresponse rate regarding household income. For instance, the 2002 MOP has a 50% nonresponse rate regarding income and in 2003 there is no data on income at all. However, from 2004 onwards the MOP nonresponse rate regarding income (3%) and mobility limitation (5%) is relatively low.

Table 8 – Missing data in trip interviews

CHARACTERISTICS	RMP	MOP 02-07	MOP 02-07 RM	MID 2002	MID 2002 RM	MID 2008	MID 2008 RM	SRV 2008
N	27.264	254.979	8.930	57.459	20.557	64.576	37.579	14.266
Distances (KM)	<0,1% (1)	<0,1% (21)	-	1,1% (613)	1,3% (276)	0,7% (443)	0,8% (297)	1,7% (240)
Duration (DAUER)	-	<0,1% (21)	-	0,2% (115)	0,2% (40)	0,1% (71)	0,1% (39)	1,4% (197)
Speed (SPEED)	<0,1% (50)	<0,1% (140)	0,1% (5)	1,6% (1.474)	3,0% (617)	1,8% (1.154)	1,9% (715)	2,2% (320)

MOP02-07RM = pooled data set (MOP-waves 2002-2007
…RM = respondents with residence in the RMP-study area only

Moreover, the data-sets contain specifications on trips with implausible distances, duration and/or speed (table 8). Missing values for distances and/or duration also results in falsifications regarding the trips' speed. The rate of reported trips which could not be included for this reason is stated in the row "speed". It is evident that the panel surveys have a high quality given their error rate of less than 0,1 %. In contrast, cross-sectional surveys have a rate of 1,9% to 3,0% in relation to the adapted basic population. Finally, additional data regarding

trips such as means of transportation employed, trip purpose as well as departure and arrival times have been collected by all surveys to a satisfying extent.

Table 9 – Item-Nonresponse in person and household data sets

CHARACTERISTICS		RMP	MOP 02-07	MOP 02-07 RM	MID 2002	MID 2002 RM	MID 2008	MID 2008 RM	SRV (2008)
PERSON DATA SETS	**N**	1.207	10.452	358	21.229	7.531	20.902	12.078	5.076
Age (ALTER)		1,5% (18)	0%	0%	0,1% (19)	0,1% (4)	0,2% (49)	0,3% (34)	0,1% (6)
Mobility restrictions (MO-BEINSCHR)		0,4% (5)	28,6% (37651; 226)	31% (1011; 10)	25,8% (5.486*)	24,2% (1822*)	24,4% (5072*; 28)	23,1% (2781*; 18)	0,4% (19)
Possession of driver's license (FSPKW)		1,4% (17)	2,8% (294)	4,6% (9)	3,4% (706*; 14)	3,4% (252*; 4)	<0,1% (12)	0,1% (9)	0,1% (3)
Nationality (STAAT1)		0%	n.a.	n.a.	26,3% (5589)	24,5% (1848)	n.a.	n.a.	n.a.
Education (SCHULAB1)		2,2% (27)	2,7% (280)	3,9% (14)	11,1% (2269*; 94)	9,9% (718*; 32)	24,5% (5072*; 39)	23,2% (2781*; 19)	1,1% (58)
Employment (BERUF1)		1,3% (16)	0,9% (93)	2,0% (7)	11% (2296*; 32)	9,7% (725*; 9)	0,1% (12)	0,1% (8)	0,4% (19)
Car availability (VMPKW)		10,1% (122)	17,2% (1.794)	12,9% (46)	30,1% (6554* ;24)	29,2% (2195*; 9)	37,8% (7901*; 3)	35,8% (4321*)	0,6% (32)
Ownership of public transport season tickets (ZEITOPNV)		7,0% (85)	6,5% (676)	5,6% (20)	26,6% (5555*; 93)	24,8% (1.835; 27)	35,8% (7467*; 27)	34% (4.089*; 11)	6,0% (267*; 37)
Car use[2] (VMPKWF)		12,8% (155)	n.a.	n.a.	26,5% (5555*; 74)	24,7% (1835*; 22)	35,8% (7467*; 13)	34,0% (4089*; 7)	0,7% (37)
Bicycle use[2] (VMRAD)		14,7% (178)	n.a.	n.a.	26,8% (5555*; 134)	25,0% (1835*; 48)	35,7% (7467*; 8)	33,9% (4089*; 4)	0,7% (36)
Use of public transport[2] (VMOEPNV)		15,9% (192)	n.a.	n.a.	26,7% (5555*; 115)	25,0% (1835*; 42)	35,7% (7467*; 10)	34,0% (4089*; 7)	0,6% (28)
Use of public transport for long distance trips[2] (VMOEPFV)		19,9% (240)	n.a.	n.a.	26,9% (5555*; 157)	25,2% (1835*; 59)	35,7% (7467*; 10)	33,9% (4089*; 6)	0,6% (31)
Normality of the day of survey (NORM)		0%	0%	0%	14,7% (3114*; 2)	14,8% (1111*)	0,1% (24*)	0,1% (15*;1)	0,3% (13)
Homogenous groups (9 groups) (PERGRUP9)		27,7% (334)	45,8% (4.788)	43,6% (156)	14,6% (3.096)	13,1% (988)	0,1% (19)	0,1% (12)	28,9% (1.468)
Homogenous groups (12 groups) (PERGRUP12)		34,1% (411)	45,8% (4.788)	43,6% (156)	14,6% (3.096)	13,1% (988)	0,1% (19)	0,1% (12)	29,2% (1.483)
HOUSEHOLD DATA SETS	**N**	626	5.893	230	8.761	3.293	9.049	5.491	2.249
Household income (EINKO1)		5,4% (34)	30,6%[1] (1.803)	28,3% (65)	13,6% (1.191)	13,3% (437)	15,3% (1.385)	15,8% (866)	15,5% (349)

[1] im Jahr 2003 keine Angaben zum Einkommen vorhanden (N=1104)
[2] people are asked to indicate wether they use certain transport modes daily, on every second or third day, weekly, monthly or never
* Frage in Stellvertreterinterviews nicht gestellt

Comparisons of trip characteristics and descriptive statistics

Several definitions have a significant impact on total distances travelled and the number of trips per day. These include the definitions of trip and mobile person as well as the attribution of a main mean of transport and the attribution of a trip purpose. If certain trips (e.g. trips outside the survey area) are excluded from the analyses of mobility participation large differences in patterns of travel behavior may occur.

In this paper an individual who makes trips outside the home is defined as mobile person. It is important for researchers to be able to distinguish between the refusal of making trips and actual immobility. The SrV places limitations on the general definition of mobile people in comparison to the other surveys employed. Only those people which make trips within or departing from the city of Frankfurt/Main are classified as mobile. If a person has made trips outside the survey area these trips are noted down; however, no trip details are recorded. If a person was immobile on the survey day he or she is asked about the reason for this immobility. Thus, the SrV defines mobility participation as "local mobility of residents in the survey area". The rate of "pseudo-immobile" persons is accordingly high. The panel surveys make a clear distinction between mobile and immobile persons (cf. table 10). In contrast, the MiD surveys allow the identification of refusers and participants whose trip diary is not available. On the one hand, the MiD percentage of people without trip diary fell significantly from 2002 to 2008 (cf. table 10). This is most likely due to a change in survey methods. In 2008 participants were asked to exclusively provide their trip data over the phone. On the other hand, the percentage of simple trip patterns (single return trips) rose from 2002 to 2008. This may indicate that telephone surveys are conducive to respondents' simplification of the trips they report in favor of certain basic trip patterns. Additionally, the MiD limits the collection of trip data to a maximum of eight trips. If more than eight trips have been made merely their number is noted. This means that 3-4% of trips recorded lack further data (cf. table 10).

The surveys discussed here categorize trip purposes slightly different. For example, the MiD data sets contain extensive information on service and accompanied trips. This includes data on the number of passengers, on whether the respondents and their companion share the same household as well as in what kind of relation they stand to one another (e.g. mother & child, colleague, employer & employee). In contrast, the SrV lacks the category of service and accompanied trips entirely. Secondly, there are also differences in the manner the surveys handle round trips and the way home. The MOP, RMP and SrV classify approx. 40% of the trips as trips home. Significantly, trips classified as such lack a thematic reference required for analyses and modeling. Faced with this problem KLOAS ET AL. (2001) developed a method for the MOP which attributes of trips home to their initial motive. If trip chains only consist of return trips the purpose of the trip to a destination is attributed to the way back to the origin. If a sequence of trips with different purposes is reported the researchers decide according to certain criteria deduced from the characteristics of the trip chain. If data-sets are not harmonized accordingly substantial differences in the attribution of trip and purpose as well as in the distances traveled are likely to result (comp. table 10).

If one aims to compare or combine data-sets for analyses it is necessary to consider the possibility that a range of different survey methods have been employed to collect data on the number of trips and chosen transportation means. Finally, the use of the SrV data-set

requires a considerable amount of preparation work. Particularly, the strong focus on local travel limits the comparability of SrV data.

Table 10 – Descriptive statistics on different trip characteristics

CHARACTERISTICS	RMP	MOP 05-07	MOP 05-07 RM	MID 2002	MID 2002 RM	MID 2008	MID 2008 RM	SRV 2008
N (Number of days)	8.449	33.943	1.379	21.229	7.531	20.902	12.078	5.076
MOBILITY PARTICIPATION ON DAY OF SURVEY (Total daily number of Trips)*								
0 Trips	8,1% (682)	8,5% (2.869)	8,8% (122)	11,5% (2.440)	11,6% (873)	9,7% (2.037)	10,5% (1.273)	10,0% (509)
1 Trip	5,0% (424)	3,3% (1.105)	3.3% (46)	4,7% (993)	4,4% (330)	7,1% (1.491)	5,0% (424)	1,1% (58)
2 Trips	34,1% (2.883)	30,2% (10.265)	26,8% (369)	23,6% (5.004)	23,0% (1.732)	25,2% (5.265)	34,1% (2.883)	35,3% (1.794)
3 Trips	11,6% (983)	11,4% (3.879)	10,3% (142)	9,5% (2.012)	9,6% (724)	11,4% (2.382)	11,7% (1.415)	8,7% (440)
4 Trips	21,5% (1.814)	21,7% (7.367)	23,2% (320)	16,8% (3.574)	16,8% (1.263)	17,6% (3.686)	17,8% (2.150)	19,5% (988)
5 Trips	6,6% (561)	7,9% (2.687)	8,8% (122)	6,7% (1.425)	7,4% (555)	7,8% (1.638)	6,6% (561)	5,7% (288)
6 Trips	6,8% (571)	8,2% (2.767)	9,6% (133)	6,7% (1.432)	6,5% (490)	7,3% (1.526)	6,8% (571)	5,8% (292)
7 Trips	2,3% (197)	3,5% (1.191)	3,1% (43)	2,7% (583)	2,8% (208)	3,4% (714)	3,6% (433)	1,9% (98)
8 Trips	4,0% (334)	5,3% (1.813)	5,9% (82)	3,7% (791)	3,9% (296)	4,6% (953)	4,3% (518)	3,2% (163)
Not applicable	0%	0%	0%	1,1% (226)	1,2% (88)	1,3% (266)	0%	2,4% (122)
Trip diary not available	0%	0%	0%	12,9% (2.749)	12,9% (972)	4,5% (944)	4,6% (560)	6,4%[1] (324)
TRIP PURPOSES*								
N (Number of trips)	27.263	117.922	4.936	57.459	20.557	64.576	27.263	14.260
Work, education	16% (4.478)	16% (18.438)	14,6% (722)	14% (7.768)	13,0% (2.665)	18% (11.787)	17,5% (6.585)	27% (3.884)
Leisure	11% (3.074)	19% (21.816)	18,4% (910)	21% (12.158)	20,9% (4.303)	27% (17.319)	27,1% (10.178)	18% (2.609)
Shopping	11% (3.070)	17% (20.312)	18,3% (904)	12% (7131)	13,4% (2.759)	18% (11.451)	18,7% (7.038)	11% (1.576)
Service and accompanied trips	15% (4.029)	6% (6.659)	5,7% (280)	14% (7.949)	13,9% (2.867)	17% (11.157)	16,7% (6.281)	0%
Way home or other	46% (12.612)	43% (50.697)	42,9% (2.120)	39% (22.453)	38,7% (7.963)	20% (12.862)	19,9 (7.497)	43% (6.191)
DISTANCES PER PERSON PER DAY ACCORDING TO TRIP PURPOSES*								
N (Number of days)	8.449	33.943	1.379	18.254	6.471	19.692	11.353	4.954
Work, education (km)	10,4	9,5	12,8	7,4	5,9	9,2	8,3	7,7
Leisure (km)	4,2	8,0	8,9	8,3	8,6	9,4	9,8	4,1
Shopping (km)	1,9	3,5	4,1	2,2	2,1	2,7	2,5	1,0
Service, accompanied trips (km)	1,4	1,6	1,2	1,4	1,3	1,7	1,4	-
Way home or other (km)	22,9	19,5	20,8	14,1	14,9	11,7	11,4	11,3

* unweighted Values
[1] According to the SrV focus on domestic travel, "trip diary not available" in SrV means that people were mobile on the day of survey but outside of Frankfurt/Main

Table 11 – Indicators for travel behaviour and modal split

CHARACTERISTICS	RMP	MOP 05-07	MOP 05-07 RM	MID 2002	MID 2002 RM	MID 2008	MID 2008 RM	SRV 2008
Trips per mobile person per day (Mean)	3,5	3,8	3,9	3,6	3,7	3,7	3,7	3,5
Distances per mobile person per day (km) (Mean)	44,4	45,9	52,3	40,8	37,9	40,0	38,2	27,2
Trip duration per person per day (min) (Mean)	85,6	90,3	96,0	84,1	87,2	84,1	86,2	70,3
Average trip distances (km)[1]	12,6	12,1	13,3	11,4	10,5	11,9	11,5	8,5
Average trip duration (min)[2]	24,1	23,6	24,0	22,9	23,5	23,8	24,4	21,9
MODAL SPLIT								
Walk	18,6%	20,5%	21,2%	24,0%	25,9%	25,7%	26,8%	31,4%
Bicycle	7,6%	10,9%	8,9%	5,2%	7,2%	7,9%	10,0%	13,6%
Moped, motorbike	0,6%	1,0%	0,1%	0,8%	0,8%	0,6%	0,6%	0,5%
Car (driver)	50,5%	45,2%	48,6%	42,5%	38,0%	41,1%	37,1%	25,7%
Car (passenger)	12,3%	12,8%	11,9%	19,1%	17,4%	15,3%	14,2%	8,0%
Public transport	8,1%	7,9%	8,0%	6,8%	9,2%	8,2%	10,2%	20,2%
Long distance travel (Train, air-plane, other)	2,0%	1,3%	0,9%	0,4%	0,4%	0,5%	0,5%	0,3%
Other	0,2%	0,4%	0,2%	1,0%	0,8%	0,8%	0,7%	0,3%
Not applicable	0%	0%	0%	0,2%	0,2%	0%	0%	0%

1 Trips with a length of more than 1000km excluded
2 Trips with a duration of more than 600min excluded

Conclusion and next steps

In future, the development of transport demand will show considerable differences at the regional level. This concerns the direction as well as the dynamics of transport demand. In light of this heterogeneity the demand for more diverse regional data sets increases. If nation-wide valid indicators of travel behavior are used for regional planning, transport modeling and forecasts substantial miscalculations can occur. Moreover, the complexity of research issues and transport models is continually increasing. Presently existing data-sets from national household surveys are no longer able to provide all of the required information. Therefore, continuously conducted, regional surveys regarding travel behavior are a significant resource. Additionally, the capacity to combine already existing data-sets is becoming increasingly important.

From 2007 onwards the Rhine-Main-Panel provides data on travel behavior for the region of Frankfurt Rhein-Main. The RMP is a reliable and thorough basis for the analysis of mobility

behavior in this region. The RMP allows analyses of the mobility in the course of the week. Additionally, cause-effect-relations can be examined (e.g. analysis of context changes, effects of policy measures). If the data of several waves is available spatial disaggregated analyses become possible. These analyses are enabled by the aggregation of certain spatial information within the data sets.

This paper presents a thorough analysis of the RMP, MOP, MiD and SrV surveys in terms of the survey design and the data collected. The surveys discussed in this paper contain a range of differences which need to be taken into account if one aims to compare and combine their data sets. The following survey characteristics have to be considered: the basic population and the reference for the basic population, the survey area, the survey method, the time period in which the survey is conducted, the sampling method, the net-sample definition of a complete household, the percentage of proxy interviews, the use of incentives to motivate participation and the design of the survey documents. Data-sets have been combined in consideration of all the differences stated above. These combined data-sets contain all the surveys' comparable information. Moreover, recommendations have been made in terms of possible applications. Nevertheless, the differences detected in this paper need be payed consideration in the work that follows.

Our next step will be to examine the extent to which RMP data can be enhanced by additional spatial and structural information. In particular, we will evaluate whether such enriched data-sets enable the analysis of multimodal mobility patterns. Moreover, we will conduct a more in depth analysis of the impact the methodological differences stated above have had on the data collected. In conclusion, generalizations regarding the comparability of data-sets should be made with caution. Depending on the individual research question the differences detected in this paper need to be taken into account and the results of analyses need to be scrutinized accordingly. First results from these next steps and analyses will be presented at the conference.

Acknowledgement: This paper builds on a report written for the ivm GmbH (“Integriertes Verkehrs- und Mobilitätsmanagement Region Frankfurt RheinMain”). The authors gratefully acknowledge the support of the ivm GmbH.

References

Bihn, Friedhelm (2002): Zwei Fliegen mit einer Klappe - Haushaltsbefragungen für integrierte städtische Verkehrsplanungen (ÖV und IV). Paper presented at Workshop "Mobilitätserhebungen zum Stadtverkehr", City

BIK Aschpurgis + Behrens GmbH (2001): BIK Regionen. Methodenbeschreibung zur Aktualisierung 2000. Hamburg

Blase, J.H. (1980): Hysteresis in travel demand. In: Transportation Planning and Technology 6 (2), S. 109-115

Bonnel, Patrick & Michel Le Nir (1998): The quality of survey data: Telephone versus face-to-face interviews. In: Transportation 25, S. 147-167

Davies, R. B. & A. R. Pickles (1985): Longitudinal versus cross-sectional methods for behavioural research: a first-round knockout. In: Environment and Planning A 17 (10), S. 1315-1329

Deutsches Institut für Wirtschaftsforschung (DIW) & Institut für angewandte Sozialwissenschaft (infas) (2003): Mobilität in Deutschland 2002. Kontinuierliche Erhebung zum Verkehrsverhalten. Endbericht. Projekt-Nr. 70.0681/2001. Berlin, Bonn

Goodwin, Phil B. (1977): Habit and hysteresis in mode choice. In: Urban studies (14), S. 95-98

Goodwin, Phil B. (1997): Have panel surveys told us anything new? In: GOLOB, THOMAS F., RYUICHI KITAMURA & LYN LONG (Hrsg.): Panels for transportation planing. Methods and applications. Boston, Dordrecht, London, S. 79-96 (= Transportation Research, Economics and Policy
5)

Hensher, David A. (1987): Issues in the pre-analysis of panel data. In: Transportation Research Part A: General 21 (4-5), S. 265-285

Kagerbauer, Martin & Wilko Manz (2009): Anforderungen an Mobilitätsdaten aufgrund heterogener Entwicklung der Verkehrsnachfrage. Erfahrungsbericht aus dem Deutschen Mobilitätspanel und den regionalen Panelerhebungen. In: INSTITUT FÜR VERKEHRSWESEN (Hrsg.): Mobiles Leben. Festschrift für Prof. Dr.-Ing. Dirk Zumkeller. Karlsruhe, S. 85-101

Kitamura, Ryuichi (1990): Panel analysis in transportation planning: An overview. In: Transportation Research Part A 24 (6), S. 401-415

Kloas, Jutta, Hartmut Kuhfeld & Uwe Kunert (2001): Dynamik des Verkehrsverhaltens im Jahresvergleich. Analyse des deutschen Mobilitätspanels 1994 bis 1999. Endbericht. Projektnummer 70.557/98. Berlin

Madre, Jean-Loup, Juan de Dios Ortúzar & Jimmy Armoogum (2006): Cross section versus continuous mobility surveys: International experiences. Paper presented at 11th International Conference on Travel Behaviour Research 16.-20.08.2006, City

Murakami, Elaine, Stephen Greaves & Tomás Ruiz (2006): Moving panel surveys from concept to implementation. In: STOPHER, PETER R. & CHERYL STECHER (Hrsg.): Travel survey methods. Quality and future directions. Amsterdam, S. 399-412

Raimond, T. (2008): Moving towards continuous collection of large scale mobility surveys: are there compelling reasons? Paper presented at 8th International Conference on Survey Methods in Transport, City

Rümenapp, Jens & Imke Steinmeyer (2005): Auswertung der Erhebung "Mobilität in Deutschland" (MiD) in Bezug auf Wochen- und Jahresgang. Schlussbericht zum Forschungsvorhaben FE-Nr. 70.0755/2004. Hamburg

Schönduwe, Robert & Martin Lanzendorf (2010): Nutzung regionaler Mobilitätsdaten. Möglichkeiten zur Kombination und Harmonisierung der regionalen Mobilitätsdaten des Rhein-Main-Panels mit anderen Mobilitäts- und Strukturdaten. Unveröffentlichter Projektbericht im Auftrag der Gesellschaft für integriertes Verkehrs- und Mobilitätsmanagement mbH - ivm. Frankfurt am Main

Stenger, Horst (1994): Anforderungen an eine repräsentative Stichprobe aus der Sicht des Statistikers. In: GABLER, SIEGFRIED, JÜRGEN H.P. HOFFMEYER-ZLOTNIK & DAGMAR KREBS (Hrsg.): Gewichtung in der Umfragepraxis. Opladen, S. 42-45

Wittwer, Ricco (2008): Raumstrukturelle Einflüsse auf das Verkehrsverhalten. Nutzbarkeit der Ergebnisse großräumiger und lokaler Haushaltsbefragungen für makroskopische Verkehrsplanungsmodelle. Dissertation. Dresden

Zumkeller, Dirk, Bastian Chlond & Martin Kagerbauer (2008a): Regional panels against the background of the German Mobility Panel - An integrated approach. Paper presented at 8th International Conference on Survey Methods in Transport, 25. - 31.Mai, City

Zumkeller, Dirk, Bastian Chlond, Tobias Kuhnimhof & Wilko Manz (2002): Selektivität des Mobilitätspanels. Schlussbericht. FE-Nr. 96.0732/2002. Karlsruhe

Zumkeller, Dirk, Bastian Chlond, Peter Ottmann, Martin Kagerbauer, et al. (2008b): Panelauswertung 2007. Deutsches Mobilitätspanel (MOP) - Wissenschaftliche Begleitung und erste Auswertungen. Zwischenbericht. Erhebungswellen zur Alltagsmobilität (Herbst 2007) sowie zu Fahrleistungen und Treibstoffverbräuchen (Frühjahr 2008). Karlsruhe

Zumkeller, Dirk, Oliver Lipps & Bastian Chlond (1997): The German Mobility Panel: Options, Limitations and the complementary use of secondary data. Paper presented at International Conference on Transport Survey Quality and Innovation: Transport Surveys: Raising the Standard. 24.-30.05.1997, City

Zumkeller, Dirk, Jean-Loup Madre, Bastian Chlond & Jimmy Armoogum (2006): Panel surveys. In: STOPHER, PETER R. & CHERYL STECHER (Hrsg.): Travel survey methods. Quality and future directions. Amsterdam, S. 363-398