Workshop Synthesis: How to Collect and Use Longitudinal Data?

Bastian Chlonda*, Lisa Eckea, Mathijs de Haasb

aKarlsruhe Institute of Technology, Kaiserstr. 12, Karlsruhe 76131, Germany
bKiM Netherlands Institute for Transport Policy Analysis, Bezuidenhoutseweg 20, 2594AV The Hague, The Netherlands

Abstract

This paper is based on the presentations, papers and discussions of workshop “Longitudinal data” at the ISCTSC conference in Porto Novo in March 2022. The presentations and discussions are primarily based on three different panel surveys that represent different approaches and designations. All these surveys have to struggle with typical challenges. The paper discusses the analytical characteristics and the designation of longitudinal surveys, addresses the problems and challenges of longitudinal surveys and will sum up the discussed strategies, in which a long-term continuity of such surveys can be secured.

© 2023 The Authors. Published by ELSEVIER B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the International Steering Committee for Transport Survey Conferences (ISCTSC)

Keywords: Panel survey, longitudinal data, travel, ISCTSC

1. Why longitudinal surveys?

Travel demand is influenced by various external processes and disruptions, e.g. the impact of the COVID-19-pandemic. The outcome of such sudden and unexpected events are superimposed by technological, economic and demographic processes, changes in the world of labor (e.g. by reasons of homeoffice) etc. with impacts on travel behavior. Furthermore, these developments are interwoven with different interventions and implemented measures in the transport system. We assume that the observed behavior results from all these beforementioned factors and processes. Altogether we would like to describe travel behavior and to give answers “why” something changes according to the central idea of travel behavior research; how and why people travel and how and why travel behavior changes.

Data are necessary to measure the same individuals (or at least the same populations) in different situational contexts to understand people's behavior. Based on these considerations, travel behavior must be observed from a...
“longitudinal perspective”. I.e., a snapshot-oriented survey (i.e. the travel of just one day only once in a while) does not allow for the detection and much less for the identification of the underlying mechanisms of changes.

Against this background, data and information are necessary to enable the researchers to understand which ongoing processes are occurring and how changes happen. This underlines the need for the temporal i.e. longitudinal dimension. However, to include the temporal resp. longitudinal dimensions in our survey, we have to consider different aspects.

- The temporal dimensions can be brought in by the same or different individuals.
- The rhythms of travel behavior during consecutive days and changes occurring between successive periods have to be regarded as different dimensions.
- And as mentioned above, we measure combined changes. These are caused by both external factors and endogenous interventions and additionally – very likely - also by methodological changes. I.e., the survey outcome can be blurred by methodological effects and artifacts.

Against this description, the workshop How to Collect and Use Longitudinal Data? at ISCTSC 2022 aimed to discuss the characteristics but also methodological challenges of different types of longitudinal surveys. The paper is intended to give an overview of the current discussion state and does presumably not cover all relevant aspects. The paper ends with some hints which can be helpful when implementing (or updating) a longitudinal survey. The open questions can be understood as hints for further research.

2. Definitions, characteristics and applicability of longitudinal data

We have to start with some definitions and classifications based on the definitions of former ISCTSC conferences (Zumkeller et al., 2006).

Repeated surveys within a population at different points in time with different individuals (completely independent samples) are a series of cross-section surveys. They can be used for temporal analyses by means of age-period-cohort-models (APC) and treated in so-called pseudo-panels (as not being a real panel as defined further below). This follows the idea that certain cohorts (as “generations”, defined by the period of birth) are followed longitudinally over time and that they age with comparable biographies, attitudes and experiences. The amount of information available is restricted, as only the “net” changes between any points in time can be analyzed (which is possible by the comparisons of cross-section surveys). The presentation of Myhrmann. et al. 2022 illustrated the applicability of this kind of longitudinal data in a pseudo-panel approach.

This format of surveys (repeated cross-sections) follows the tradition of travel data collection for the provision of statistics. However, for identifying and assessing changes, more about processes must become known. This is a challenge since changes can occur in various directions, but mostly the changes do not counterbalance each other (e.g. trends). It is insufficient to look only at the “net changes” to understand these trends and their determining factors.

Panel surveys allow for intraindividual observations and a longitudinal perspective. Individuals are surveyed at different points or over different periods of time (longitudinal perspective), which is the idea of a panel survey in its original sense. As we have changes in different directions, we cannot only consider the net changes of the aggregated behavior but much more the “gross changes” as the changes of every individual within the sample.

We know by former research that for measuring and explaining changes in travel behavior” a short observation period is insufficient as the intraindividual day-to-day variability is comparably high (see e.g. MOBIDRIVE (Axhausen et al. 2000, or the Upsalla panel in the 80’s (Hanson, Huff 1988). Here we come to another form of a longitudinal observation or longitudinal perspective, i.e. the multiday survey as a continuous observation. Both the German Mobility Panel (MOP: observation period of 7 days) as well as the Netherlands Mobility Panel (MPN: observation period of 3 days) follow that idea (De Haas et al., 2022a).
Generally spoken, we have three different forms of surveys to integrate the longitudinal perspective:

1. Following the same individuals over time by consecutive waves, asking them the same questions (a panel in its closer definition).
2. Doing the same surveys with independent samples (series of cross sections surveys)
3. Observing people continuously over some time (which is a special case for a panel).

It should be mentioned that 1. and 3. are typically used in combination. This follows the idea that willing-to-cooperate participants at hand allows for more in-depth analysis (Haas, de M. et al. 2022).

Besides the potential to identify changes and to analyze processes, the availability of a sample of willing-to-cooperate participants for the cases of a "panel" (1.) has another significant advantage. The example of the COVID-19 pandemic suddenly brought up an external impact that needed to be observed (Ton et al., 2022). The availability of a sample with known characteristics and a known travel behavior from earlier waves allowed us to ask specific questions without a complex time consuming and costly sampling and thus enlarged our understanding process (Ecke et al., 2022b; Faber et al., 2022).

Moreover, the characteristics of panel surveys as well as the continuous observations allow for a data quality evaluation by the potential to identify the bad risks in the survey and to correct any errors and reporting flaws based on the existing context information. This additional advantage must be stressed as three of the papers presented in the workshop addressed these issues and made use of the information included in the longitudinal data (Ecke et al., 2022a; De Haas et al., 2022b; Ecke et al., 2022c).

Last not least the central aspect: Against the background of the idea, that we want to use the longitudinal data to analyse processes all these surveys should measure the same characteristics and should collect and process the data in a similar way. This has to be regarded as the central challenge which was discussed in detail in the workshop.

In the following we relate on both kind of panel surveys and also on repeated cross-section surveys as some aspects are relevant for all these types of surveys. The main issue of the discussion has been how sensitive surveys are towards any changes. Therefore, all of these surveys need a methodology that identifies travel behaviour changes that are not disturbed by methodological artifacts.

3. Challenges for longitudinally oriented surveys

The collection, use and interpretation of longitudinal data results in typical challenges, which are more or less relevant for all “longitudinally” oriented surveys. These challenges are often closely related to each other. However, the solutions to overcome these challenges produce dilemmas and must be balanced about their advantages and drawbacks.

3.1. The challenge of continuity-paradigm

We defined continuity as a basically unchanged design, meaning we have no methodological or other changes that might hide or mask any observed behavior or its changes. This means: The continuity requires a comparability of the results of all survey waves, independently if these are repeated cross-sections or a panel. This implies that the same approach in terms of the methodology, the recruitment procedures, the treatment and compilation of the data (cleaning, weighting, treatment of missing values) etc. must be secured for all panel waves, resp. the repeated cross-sections. However, we are aware that this is difficult to achieve at least against the background of surveys performed repeatedly for many years as obviously researchers and market research firms are confronted with massive changes in terms of survey methodologies on the one side and the acceptance of surveys on the other. This leads to a dilemma. As the data must be comparable and free of methodological artifacts which might be a result of innovation, this “continuity-paradigm” prevents any innovations in a repeated survey.
3.2. The challenge of aiming at representativeness and the selective participation

We know that installing a panel is a challenge to be “representative” for a long time. The reasons for this are person-group specific selectivity resulting from the response burden and the dropouts between waves (“attrition”). Many efforts are necessary to make a panel survey viable at all. Compromises are essential and the recruitment will depend on the designation of a survey. However, one must be aware that having a representative sample is also not always necessary. The survey should be appropriately and adequately designed for the research objective.

- The German Mobility Panel (MOP) aims to measure the travel behavior in Germany on time series. To keep the sample “representative” the participants are kept for a maximum of 3 years in the sample and each year a refreshment has to take place (“rotating panel”). This approach is costly (each year a new cohort must be recruited), however, must be understood against the designation of the MOP as a multipurpose instrument, which also has to provide each year statistically sound figures.

- The Netherlands Mobility Panel (MPN) approach is different as participants are recruited using a consumer open access panel, which only addresses the online population and is thus biased. However, with yearly and continuously performed national household travel surveys (NHTS) (e.g. ODiN) at hand in the Netherlands this compromise is a sensemaking approach, as the main goal of the MPN is to study underlying mechanisms of travel behavior change and not to provide statistical figures.

- The panel of Dutch Railways (Ton et al., 2022) also aims to understand processes within their ridership. The uncontrolled recruitment of customers has to be regarded as not representative, but allows for observing and analysing the processes.

The main challenge within this context in the last years has been the decline of participation rates. Reasons are seen by respondent burden (too many or too long surveys) and data privacy concerns. This is obviously a worldwide phenomenon. This mainly and stronger affects young adults (shown in all of the surveys discussed) and became obvious in comparing the MOP and MPN (De Haas et al., 2022a).

We are aware that the more complex a survey is, the more difficult to find participants at all and the more selective and the less representative the sample. This is especially the case for panel surveys because of the repeated treatment resulting in a high respondent burden. We, as researchers, try to react and install adapted forms of recruitment as well as sampling procedures. These result in different representations of the basic populations, i.e. the population in the different survey waves, resp. cohorts in rotating panels. These samples are likely to be different and not directly comparable. I.e. any measures and recruitment design changes will end up in different samples with – likely – a different behavior. The recruitment should aim at people who will stay for some time in the survey to allow for the analyses of change. Hard-to-reach groups can be incentivized, however, this is likely to also result in some form of selectivity which must be controlled.

Here a compromise must be found, which has to be balanced between the representativeness on the one side and on the other to find participants at all - independently by which approach recruited. The determination of a best practice takes place in a field of tension between limiting costs, contractual agreements and the demand for a sample that optimally fits to the certain needs of the study.

3.3. The respondent burden - conditioning, attrition and fatigue

Panel surveys - both repeated and continuous observations – query the same to the same individuals. This is very likely to be felt as “boring” for the participants. As a result, they may lose interest in participation, which endangers the idea of the panel survey. And individuals learn to adapt their reporting behavior to ease their burden. However, this produces artifacts. Methods exist to identify and quantify these effects (Chlond et al. 2012). Several papers and contributions related to the workshop dealt with these aspects (Ecke et al., 2022a; De Haas et al., 2022b) and therefore cannot be completely ignored. Conditioning means that participants learn to report “smarter” (by e.g. bunching trips together to ease their burden). This “conditioning” affects the comparability of the results between years. However, it
should also be noted that especially in times of the COVID-19 pandemic, the observation may be biased because travel behavior changes due to external influences (Ecke et al., 2022b; Faber et al., 2022; Ton et al., 2022).

Attrition means that participants drop out between waves. While these dropouts may positively affect the quality of data as the dropouts are likely to report less accurate even before (Chlond et al. 2013) managing attrition is crucial for a panel. As all these effects are an outcome of respondent burden. The burden must be kept reasonable and people must be kept motivated (e.g. by an incentive). The latter is likely to affect the representativeness again.

A thoroughly designed panel care can help, e.g. regularly keeping in contact with participants, offering incentives e.g. by presenting the interested participants results of the survey. This is likely to affect representativeness (“selective participation”), resulting in new dilemmas.

3.4. The challenge of new survey technologies and methodologies

We aim at the “ground truth”, i.e., we want to have unbiased data. In the nature of research, we are likely to improve our understanding of methodologies and effects. However, this is also likely to impact the comparability of the results of any previous study if we change anything in the methods. As a consequence, the interpretation of any changes becomes difficult. This does affect not only panels but also the repeated cross-sections with independent samples.

Against the challenges of respondent burden e.g. technological means (GPS-based / smartphone-based surveys) make sense and would be at least more efficient, especially in the case of a multiday surveys. This kind of “passive” data collection eases the respondent burden. However, this kind of data is confronted with privacy concerns and we still miss central information, e.g. for the case the travel behavior is completely reported, the context information is missing (e.g. cell-phone data without a socio-economic understanding, we have no trip purposes.)

It became clear in the discussion that switching to another mode of data collection will produce completely different results. An issue of discussion was, how far automatic data collection methods can clearly distinguish relevant trips or stages. They can detect and distinguish modes to a certain extent, however, the question of when a trip starts or ends cannot be answered. And we are aware that we will not address complete populations by means of this approach (e.g. senior citizens are more likely to be missing). However, this might be less relevant as no significant data about total travel volumes are required. Some potential to integrate gamification to overcome some of the issues is seen, e.g., allowing the participants to control the output and overcome and at least lessen the privacy concerns. From the 2021 perspective, passive data collection is still not yet able to substitute conventional diary-oriented surveys. However, passive data collection is regarded as a substitute for implementing and designing completely new surveys. Furthermore, such methods could possibly meet the needs of young adults for participation, who are first challenging to recruit and second hard to keep in a panel. This is another issue of a necessary balancing. We have always to balance the advantages against the drawbacks of any solutions.

And beyond, our pursuit to find the “ground truth” is likely to be unsuccessful as we are a little bit behind what is going on in new approaches - not only in the research “market”.

3.5. The role of survey frequencies

Typically, NHTS aim to give an overview of the current situation of travel behavior and are performed every 6 – 10 years. During this period, many things are likely to change. As these surveys - this is an issue of the repeated cross sections only - are only performed “seldom” they are more likely affected by methodological adjustments and the changed survey market environment. This results in data that do not allow the separation of methodological from behavioral effects and changes. Obviously, it is necessary to perform such surveys frequently and continuously as it is done e.g. in the Netherlands with the OViN survey (Statistics Netherlands 2022). In the latter case, a methodological change can be assessed. The challenge is in the administrative implementation of the regular funding for such an approach.
3.6. Conclusion so far - Dilemma Circles

Altogether we are confronted with different dilemma circles. We have problems, and our remedies and solutions will likely create new challenges and problems, which all interrelate. Any methods to improve the data will result in collateral effects affecting quality and the continuity.

4. Consequences for longitudinal surveys with a long-term orientation and what to think about when updating an existing or installing a new longitudinal survey?

In most cases, the survey design is chosen at the time of the first implementation of the survey. Therefore, the longest-lasting surveys with more or less unchanged designs are based on a methodology of the implementation period; i.e. what was the best option at that time against the background of the intended purpose of a survey†?

To keep continuity, fundamental design changes become difficult. The risk of methodological artifacts can be interpreted as behavioral changes in the worst case. And this makes a design change unlikely and are therefore not wishful. It becomes a more serious problem when many new approaches become available, which sound promising as remedies within a short time.

Therefore regularly performed NHTS designed for large cross-section samples also aim at replications in the design (as far as possible) to allow for comparisons and data as time series.

4.1. Consequences and conclusions for updating an existing survey

Advances in research, technology, and survey methodologies are likely. We want to compensate the declining willingness to participate by adapted and new forms of recruitment, approaches to ease respondent burden, keep cost in a reasonable frame, etc.. This entices us to apply any upcoming new ideas and solutions. However, all modifications have to be handled with care against the background of the overall objective and the paradigm of continuity:

1. All changes to be implemented must be proven in other applications and must have been tested in advance before the implementation in the survey. This is also the case for more or less seldom performed survey waves or repeated surveys with low frequencies (i.e. performed less than yearly).

2. All methodological changes should be “looped in” slowly and with care. Looping in means that at the beginning, only a portion of the sample will be treated with the new approach, the consequences have to be tested and assessed, and successively the new approach can replace the old one. This mainly affects continuously performed surveys or highly frequent surveys. The existing approach and the new one have to run in parallel. The consequences for the results and the person groups represented in the different methodologies must be assessed with care. The new design can then replace the old one, however, the consequences and the impact on the outcome must be analyzed, evaluated and well-documented. This can be the case e.g. even a changed weighting procedure. A modified weighting approach is comparably easy to install as both weights - the new and the old ones – can be provided with the relevant information. Such an approach requires a reweighting of the old data (i.e. a new reweighting of former waves in the past as well – which in some cases (missing information in the data) is impossible.

---

† Since 1994, the German Mobility Panel is still based on a paper based approach (PAPI), resulting in declining participation rates especially among young people and by reasons of the recruitment process (quota sampling by RDD). This was state-of-the-art at by the start of the survey. However, it is likely to be costly today. The Netherlands Mobility Panel introduced in 2013 is based on the participants of an access panel (“consumer panel”) and collects the information by CAWI only.
3. Too many changes in the methodology at the same time are difficult to handle. The different effects cannot be controlled and separated. Therefore, only changes should be implemented for which the methodological effects are well known (or can be assessed in terms of their relative impacts).

4. Transparency: All changes must be documented, e.g., in the meta-information and the coding plans of the survey, to allow future researchers to use the data.

The alternative for a looping-in approach could be a sharp cut. A completely new design (which should be thoroughly tested before implementation) will be introduced. However, this could result in an incomparability that is not wishful against the idea of a longitudinal survey aiming at continuity. Consequently, the question was discussed when it is the right time to make such a cut. The outcome of the discussion was “never”. We aim after the COVID-19-pandemic to capture and identify a “new normality”. However, we will only know years later when this happened (and if at all), and we are likely to choose the wrong point in time for a change in design. This again underlines the need to loop in innovations to assess their effects.

4.2. Consequences for the installation of new longitudinal survey with a long-lasting orientation:

Longitudinally oriented surveys are usually intended to provide data to analyze processes for a long time (years). The survey must therefore be designed for a long-lasting perspective. This is in contrast to advances (technologically or methodologically) and the challenges of the declining and selective willingness to participate. This latter aspect mainly affects panel surveys, where we repeatedly ask the same respondents the same things, which affects fatigue and panel conditioning. Against this background, we are in a dilemma from the early beginning when planning and implementing a new survey:

1. The intended use of a panel has to be defined at the early beginning. This affects the general design, the recruitment, the representativeness and of course cost. It must be decided if a panel will rather aim to explain changes and will not completely represent all population subgroups or if it is also designed to provide statistically sound figures. In the latter case, such a multipurpose approach affects many of the following aspects.

2. A design must be chosen which is (at the time of the implementation) regarded as future-proof. This will likely include technology and methodological innovations, for which the applicability has already been proven.

3. The design must be tested thoroughly before an implementation not to be forced to make any alterations soon after the implementation.

4. An organizational structure must be defined, which allows for the continuity of the survey (contractors, buyer, market research firm) and should provide fallback options, e.g. if one of the contractors withdraws.

5. A legal structure must be found and defined which does not run into restrictions (e.g. restriction of the duration for which participants can be recruited, storing addresses of participants etc. against data protection legislation etc.). This is narrowly connected with the organizational structure.

6. The organization of the survey must be made robust also against administrative hurdles. If privacy legislation (GDPR), as well as long-lasting funding, becomes difficult, an organizational approach must be chosen that also allows for continuity (from the side of the contracting entities as well as the contractor).

7. Panels result in a relatively high respondent burden and long-lasting participation is crucial. Any methods to ease participation are welcome. This can be done by integrating technological approaches (mainly GPS-based smartphone surveys) with a high degree of automatization.

8. The technological-based survey will result in certain selectivities. Against the idea of having a panel that aims to analyze changes, a certain amount of selectivity is likely and must be accepted. However, it must be understood as the lesser of two evils.
9. Against the background of the selectivities and the resulting high cost, approaches will become more accepted, which potentially do not represent populations completely:
   a. Accept limited representativeness and a certain amount of selectivity.
   b. Combine for longitudinal observations a panel with a continuous and repeated cross-section observation: For the case of a (selective) panel, a regularly performed cross-section survey as NHTS allows to identify how far the panel represents the population and its behavior. In turn, the panel allows for the interpretation of the cross-section results.

10. Plan the cross-section-surveys regularly not to endanger their applicability (i.e. a survey every 5 years allow the breakdown in each five years generation etc.). Ideally, the cross-section survey will be run continuously.

5. Conclusion

Against the background of the discussion, research priorities have been identified as follows:

1. The application of smartphone data to generate individual trip data can be achieved by actively integrating participants in passive data collection (gamification, allowing people to control what the smartphone records). How far do these surveys endanger acceptance on the one hand and how far this helps to overcome the existing obstacles?

2. The transferability of survey data generated by different methodological approaches: Discussion of experiences made with design changes resp. changed survey methodologies: Is it possible to combine or fuse data generated by different methodological approaches reasonably? What does this mean for mixed-mode surveys (intra- and interpersonally)?

3. Sample representativeness: Who is represented in different samples, who is not and why? How much is the resulting error (missing data). How much representativeness is needed at all? Is it necessary to have a perfect representation against cost-benefit considerations?

4. Do we need “standards” and “norms” for the implementation and execution of surveys, e.g. comparable to the ISO 20252?

**Last but not least - The role of patience**

Every longitudinal observation is like a fruit tree that grows for many years until it is fruit-bearing. And the amount and quality of the fruits will increase with every additional year resp. wave. However, the funding institutions and researchers are increasingly impatient with increasing duration. And as travel behavior only changes with low speed, the usefulness will become obvious only after several years (e.g. when trends become obvious) or exactly when anything happens. Against the background of the COVID-19 pandemic, we acknowledge the existence of ongoing continuous and longitudinal surveys that allowed for assessing the effects.

**Acknowledgements**

The results of the discussion are collaborative results of the participants (in alphabetical order):
Barbara Christian (France), Basil Schmid (Switzerland), Bastian Chlond (Germany), Jean-Loup Madre (France), Lisa Ecke (Germany), Mathijs de Haas (The Netherlands), Roel Faber (The Netherlands), Danique Ton (The Netherlands), Marcus Skyum Myhrmann (Denmark), Quin Zhang (Germany), Raky Julio (Spain), Rui Colacau, Tim Wörle (Germany)
References


Ecke, L.; Chlond, B.; Vortisch, P. (2022a): How late-reporters affect the overall data quality in longitudinal panel surveys – experiences of the German Mobility Panel


De Haas, M.; Kroesen; M.; Chorus, C; Hoogendoorn-Lanser; S.; Hoogendoorn; S. (2022b): Didn’t travel or just being lazy? An empirical study of soft-refusal in mobility diaries

Faber, R.; Hamersma, M; De Haas, M. (2022): Measuring the effects of COVID-19 on our activities, work and travel behaviour using longitudinal data with the Netherlands Mobility Panel’

Myhrmann, M. (2022): A Cycling Population in Decline: A Danish Pseudo-panel Analysis
