



MINISTERIE VAN VERKEER EN WATERSTAAT

PROJECTBUREAU INTEGRALE VERKEERS- EN VERVOERSTUDIES

Plesmanweg 1-6 (F vleugel, 3e etage), Postbus 20901, 2500 EX Den Haag.  
tel. (070) - 351 74 54. fax (070) - 351 63 64

# EVALUATIE VAN HET LONGITUDINAAL VERPLAATSINGSONDERZOEK

September 1992

Vakgroep Sociologie  
Rijksuniversiteit Utrecht  
Heidelberglaan 2  
Postbus 80.140  
3508 TC UTRECHT

P.G.M. van der Heijden  
H.F.E. Smit  
G.H.C. van Gils

ISOR Methods Series MS-92-9  
ISBN 90-5187-123-6

## Inhoudsopgave

|   |     |
|---|-----|
| Summary   | i   |
| Voorwoord   | x   |
| I. Inleiding  | 1   |
| II. Over het LVO:   | 5   |
| 2.1 Het LVO een beknopte geschiedenis                             | 6   |
| 2.2 Het LVO doelen, opzet en problemen                            | 15  |
| 2.3 Het LVO hantering van de problemen                            | 28  |
| 2.4 Het LVO vergeleken met het OVG en andere vervoerspanels       | 36  |
| III. Over de inhoud:  | 44  |
| 3.1 Benutting van de inhoudelijke studies door beleid             | 45  |
| 3.2 Inhoudelijke studies, conclusies en validiteit                | 52  |
| IV. Voor de toekomst:   | 59  |
| 4.1 De mogelijke rol van commerciële onderzoeksbureaus            | 60  |
| 4.2 Voorstellen m.b.t. toekomstig longitudinaal beleidsonderzoek  | 67  |
| V. Conclusie  | 81  |
| VII. Bijlagen   | 1-1 |
| B.1 Bijlage 1: Een beschrijving van (buitenlandse) vervoerspanels | 1-1 |
| B.2 Bijlage 2: Een beschrijving van Nederlandse panels            | 2-1 |
| B.3 Bijlage 3: <i>Personae Informantes</i>                        | 3-1 |
| Literatuurlijst   | i   |

## Summary

### 1 Introduction

The Project Bureau IVVS commissioned the Department of Empirical & Theoretical Sociology<sup>1</sup> of the University of Utrecht to conduct an evaluation of the Dutch National Mobility Panel (LVO). The LVO panel survey started in 1984 and was concluded in 1989 after ten waves. Each wave consisted of approximately 1700 households. The main purpose of the evaluation of the LVO panel survey was to look into the methodological and statistical complexities which accompanied this panel survey and how these problems have hindered the attainment of the research goals of the LVO panel survey. Therefore, in light of the above mentioned purpose, we have described the panel history, identified the main problems and described these problems. Assessments were made in order to establish the validity of the substantial conclusions which were generated by this panel survey. Some recommendations are made with regard to future design of longitudinal policy research. This summary is organized as follows:

We will begin this summary with a brief description of the Dutch National Mobility Panel (LVO) in terms of its organization, design, and some methodological problems which stemmed from the planning- and data collecting phase (2). This will be followed by a description of how the attainment of the research goals was hindered by these methodological problems and their resulting biases (3). We then give consideration to the correction procedures which were employed to correct the biases, and how this complicated the different types of analyses (4) and affected the validity of the substantive research findings (5). We will conclude this discussion by indicating how the methodological problems can be avoided, or reduced, in the field during the data collection phase and be handled during the analysis phase (6). Finally we will indicate a few directions in which future developments of longitudinal policy research can be expected to be optimal (7). The reader is advised to refer to the main report for a more thorough and detailed discussion of each of these points.

### 2 The design of the LVO

#### 1. *Organisational background*

The panel study was financed by two parties: the Project Bureau for Integrated Transport Studies (PbIVVS) and the Directorate General for Transport (DGV of the Ministry of Transport & Public Works (V&W)). In addition to these two commissioning institutes, there were other parties that had an interest in the panel study and could influence decision making, such as: the Department of Planning (RPD, of the Ministry of VROM), and the Traffic & Transportation Research Division (DVK). Representatives of the Dutch Railroads (NS) and other Public Transport Agencies were involved parties as well. At different stages of this panel study, different goals and expectations were formulated and reformulated. In this context, the Dutch National Mobility Panel was, at times, referred to as a 'multi-purpose panel'. It is not unlikely that the concept of a 'multi-purpose panel' finds its source in the presence of the many parties which were involved, each representing different interests and points of view.

---

<sup>1</sup> As from January 1992 this department became the *Department of Sociology*.

Therefore the planning of this research did not develop in a straight and smooth way. On the contrary, it was, more often than not, subjected to the whims of political and financial opportunity and creativity, which, in turn, led to pressures such as tight financial and time budgets. All this caused turbulence during the designing phase of the research. The organization of the panel was further complicated by its use of different research institutions to carry out this panel study. Creating a sample frame and drawing the sample from it, plus data-collection and data-management was conducted by the Institute for Longitudinal Policy Research (ILBO) together with Intomart (for the field work). Analyses were performed mainly by Bureau Goudappel & Coffeng (BGC) and some other research institutions and (individual) consultants.

## 2. *Panel design*

It was suggested as early as 1978 that a panel should be created to study mobility over time. Two feasibility studies were conducted, each assessing the costs and benefits of a different type of panel. In 1983 it was decided to design a panel study consisting of about 2000 households for a, then unknown, number of waves. Consideration was given to a stratified sample, refreshment of drop-outs and the questions to be asked. The panel was to serve two related goals: (1) the description of the change of the Dutch mobility patterns over time and (2) the causal analysis of the change of Dutch mobility patterns over time.

In the autumn of 1983 the political opportunity to actually launch this panel study arose when the Dutch Parliament expressed the need for an evaluation study to assess the impact of the Transit Fare Increase Plan (*Tarievenplan*) on the different segments of the Dutch population and on the use of Public Transport. This added a new goal to the panel: the evaluation of the Fare Increase Plan. This evaluation was to untangle the effects of the fare increase from other causes<sup>2</sup>. It should be stressed here that this idea rose only months before the first price increase occurred and thus little time was left to redesign the panel study in order to prepare it adequately for its new task. In March 1984 the panel was launched in a rather premature state. The before mentioned 'multi-purpose' character of this particular panel study and the turbulence around it, did naturally affect the quality of its design. We will now turn to this point.

The desired sample size of the panel study was finally set on approximately 1800 households. Within this sample there was a non-proportional oversampling of population segments which were considered to be politically interesting, such as: transit users, single parent families, low income groups. These choices were politically motivated. Also certain population segments had lower response probabilities than others and these groups were, by design, oversampled as well. The selection of municipalities was also partially influenced by political considerations, rather than by methodological rules alone. This resulted in an accidental over-representation of urban municipalities (as compared to rural municipalities). All these factors resulted in a unbalanced data-set.

In order to evaluate the Transit Fare Increase it was decided to conduct two waves of data-collection per year, one in spring and the other in autumn. During subsequent analyses this turned out to be a problem as season effects could not be distinguished from effects

---

<sup>2</sup> A decrease in petrol prices occurred soon after the first fare increase.

caused by the fare increase, given the small number of waves at that time. Other factors contributed to the complexity of the sample as well: the LVO-data were structured hierarchically: trips per person, persons per household and households per community type. As a consequence of the data-structure, observations on the level of trips and persons are correlated.

Another design feature should be mentioned also. The questionnaires were overloaded with too many questions, due to the use of Travel Diaries which covered a period of seven days; many (additional) questions about Transit use; possession of tickets and the questions about properties of cars. This resulted in *respondent burden*, in other words, the respondents became weary of the long, really too long, questionnaires. In summary: we were faced with a complex sampling scheme and an over-loaded questionnaire plus a travel diary.

### 3. *Methodological problems in the LVO*

The design of this panel resulted in four problems all of which played important roles.

1. The sample was unrepresentative of the general Dutch population due to the disproportional over- and under-representation of certain population segments in the sample. The complex structure of the sample complicated correction procedures which were applied in the data-analysis phase of the research. We will refer to this problem as the complexity of the sample, or alternatively, the generalization problem, since the complexity of the sample caused problems with generalizing the research findings to the Dutch population as a whole. This was evident when the research outcomes from the LVO were compared with those of other studies which provided consultants with a way to assess the external validity of the Dutch National Mobility Panel. Other problems (see 2, 3 and 4 below) added to the generalising problem.
2. A massive and selective initial non-response (of 62%) also occurred.
3. Selective drop-out (attrition), from one wave to the next, happened throughout the study, averaging a rate of about 20% per (odd) wave. See figure 1 below.

**Fig.1. Attrition and sample refreshment  
in the Dutch National Mobility Panel**

|     | Wave1  | Wave3  | Wave5  | Wave7  | Wave9  | Wave10 |
|-----|--------|--------|--------|--------|--------|--------|
| (1) |        | 752    | 523    | 560    | 67     | 17     |
|     |        | ↓      | ↓      | ↓      | ↓      | ↓      |
| (2) | 1764 → | 1688 → | 1849 → | 1926 → | 1749 → | 1549   |
|     | ↓      | ↓      | ↓      | ↓      | ↓      |        |
| (3) | 828    | 362    | 483    | 244    | 217    |        |

Legend: 1 = sample refreshment; 2 = number of households; 3 = attrition

4. As a result of the over-loaded questionnaires and travel diaries, respondents gradually began to report less and less trips over time. This effect occurred not only within each measurement (within one single wave less trips were reported in the later days of the week), but also between measurements (less trips were reported in later waves). These measurement errors are referred to as within- and between measurement biases.
5. Some times, for some types of analysis, the sample size was regarded as too small. The sample size was too limited when, (1) population segments and when (2) rare events were studied, and when (3) small effects were studied.

We strongly feel that problems 1, 4 and 5 are rooted in the panel design and primarily reflect the difficulties of designing a multipurpose panel in a sensitive and changing policy context. Problems 2 is a problem common to random surveys and problem 3 to panel surveys.

### 3 Reaching the research goals: obstacles

The five problems produced both positive, as well as, negative developments. On the positive side, there is now a large body of methodological knowledge (which has been used in the Netherlands, as well as, internationally <sup>3</sup>) with respect to these problems. On the negative side, these problems stood in the way of the attainment of the research goals. These problems can be described in respect to the three research goals. The recommendations voiced at the Round Table Conference <sup>4</sup> in 1987 are illuminating in this respect. We can summarize them as follows:

1. The description of the change of mobility patterns in The Netherlands over time was considered as an unrealistic goal during the Round Table Conference in 1987 because of the five problems previously mentioned. After the Round Table Conference, consultants turned towards analyzing segments of the sampled population for specific purposes. This was, however, not a solution of the generalization problem; even subpopulations can still be unrepresentative due to selection mechanisms in initial non-response and attrition. Moreover, the number of observations tends to be very small when analyzing only parts of the sample, resulting in a low accuracy of estimates.
2. The causal analysis of the change of mobility patterns in The Netherlands over time was considered to be less influenced by the problems that played such an important role on the level of descriptive statistics. The following arguments were given at that time. Using *change-scores* in causal analysis <sup>5</sup> reduces the impact of the within- and between measurement effect. Another point is that regression-coefficients ( $\beta$ 's) are quite stable and are not easily affected by design features (such as over- and undersampling). So regression-coefficients are likely to be unbiased even when the sample is not quite representative of the population. Therefore, it was concluded that while it was not safe to generalize descriptive analyses, it was safe to generalize the

---

<sup>3</sup> Through international conferences, publications and through contacts of consultants.

<sup>4</sup> The purpose of this Round Table Conference was to evaluate the development of the panel study up to that moment and as to whether its goals could still be met.

<sup>5</sup> In many causal analyses performed on the LVO-data, change scores of the dependent variable were regressed on the change scores of the independent variables.

research findings from causal analyses. This approach, however, was not without its drawbacks. For example, it was found that not very many changes could be observed in the panel. It was, therefore, decided that more waves were needed. More waves were also needed to assess long-term effects.

3. The Evaluation of the Fare Increase Plan was never a realistic goal. First of all, the sample of the LVO was too small to adequately assess the changes caused by a fare increase of about 10%. It was entirely impossible to do this on the level of population segments and retain an acceptable accuracy of the estimates. Also a more or less theoretical problem (see 2.3.7 and 4.2.4) made such an assessment unrealistic; shortly after the first fare increase, petrol prizes dropped and it became very hard, if not impossible, to untangle the different causes.

#### 4 Analysis

The forementioned problems which arose from the planning and data-collection phase, in turn, affected the analyses by complicating matters. It is our impression that the consultants made a serious effort to lessen the effects of the problems as far as possible. As far as the consequences of these problems on the analyses are concerned, we like to make the following remarks:

1. Small numbers of observations lead to unreliability in the estimates and to insufficient power to detect true but small changes. Generally, there are broad confidence intervals around parameters estimates of models <sup>6</sup>.
2. Attempts have been made to correct the sample design (*over- and undersampling* of specific population segments) and the selective initial non-response (the differential response probabilities of the different strata). The necessary data to perform these corrections -a post-stratification- came from the Mobility Behaviour Study (OVG), an ongoing repeated cross sectional study carried out by the Dutch Bureau for Statistics (CBS). The following stratification variables were used: Household Stage, Household Income and Type of Municipality of Residence. Post-stratification is a re-weighting procedure in which the distribution of LVO-units are re-weighted in such a way that they become comparable with the distribution of the OVG-units in the stratification matrix. The LVO-data are likely to have a lower quality than the OVG-data, because corrections using only three variables are prone to be only partially successful.
3. Re-weighting has also been used for correcting bias due to selective attrition. The combined weights (of points 2 and 3) occasionally grew large. Again, only the information captured in the three stratification variables were used. Other variables that predict the (probability of) dropping out could also have been used, and, in general, it is recommended to do so when the opportunity arises. Either one corrects for selective attrition by re-weighting, or by modelling the attrition in a regression-like model <sup>7</sup>.

There is an indication that re-weighting did not sufficiently correct all selectivity due to the sample scheme, selective initial non-response and selective attrition in the LVO

---

<sup>6</sup> In some reports significance levels are not mentioned. This creates difficulties in interpreting some of the research outcomes.

<sup>7</sup> This type of (Heckman-like) corrections were considered. The consultants felt, however, that reweighting would be appropriate for the main stream of analyses.

sample. It can be shown, that after re-weighting, the households that stayed within the panel are more mobile than those households that left the panel. Due to points 2 and 3 we assume that descriptive statistics are likely to be biased. However, it is very hard to say whether points 2 and 3 also affect the parameters in the regression models (see 3.2.2).

4. The within- and between wave biases are related with the tendency of respondents to report less and less trips over time. For descriptive statistics, such as means, this is a serious problem. For such statistics a correction procedure is proposed that seems to elevate the problem somewhat <sup>8</sup>, however, for most purposes the use of descriptive statistics remains suspect, and should be discouraged. Problems in the correction procedure occur when there is a relation between change of 'between wave biases' and change in background variables. Whether both types of changes are related or not is hard to say, because there are not enough data to investigate this. For causal analysis it is likely that the within and between wave biases have caused a much smaller problem <sup>9</sup>.

So far we discussed how the methodological problems affected and complicated the analyses. We have the impression that the consultants have handled these problems as carefully as possible. And as far as the analyses are concerned, we see an increase in the level of sophistication in the analysis, accumulating, for example in Meurs' Ph.D.-thesis. We like to mention three important developments.

1. The first improvement is the shift from static towards dynamic models (models that have lagged values of dependent and/or independent variables) for analysis. In hindsight it could now be assumed that the 'old' static models are misspecifications which produced biases in estimates of regression-coefficients. Using dynamic models on LVO-data is, from the historical point of view, a novelty.
2. The second improvement concerns the use of the refreshment in the analysis. In the earlier reports information was used of only those households (or persons) who had participated in all waves. In more recent reports this was improved by *wave pooling*. However, most often consultants did not take into account that observations were correlated then. This problem is also tackled in Meurs' Ph.D.-thesis with help of *random coefficient models*.
3. The third important development is the development of a (prototype) simulation model, called MIDAS. If the simulation model is valid, then simulations could considerably help to understand the dynamics of (household) behaviour and how these affect travel behaviour. Indeed, simulations should precisely be used for gaining insight in how changes are conditional on certain parameters. Since MIDAS is, for us, a 'black box' we do not know how accurate MIDAS is in terms of predicted outcomes and consequently we have no further recommendations.

In summary: the methodological problems were all recognized in the analysis phase of the LVO-study. Consultants made a great effort to correct the resulting biases. Not all the problems could be removed however, since the correction procedures were based on certain assumptions and it is unknown whether these assumptions were fulfilled. One can only

---

<sup>8</sup> Uncorrected estimates of mobility gives the impression that there are decreasing trends in mobility, while corrected estimates show that there is an increase in mobility trends over time.

<sup>9</sup> The point is that a constant 'within bias' will be canceled when using change score. There is some evidence, however, that the assumption of a constant 'within bias' is not correct.



speculate about this. Therefore, one can not assess the extent to which the estimates *remain* biased after corrections were made. On the level of substantive conclusions of causal analysis we feel that these biases will not seriously inflate the conclusions. For the interpretation it doesn't really matter whether a coefficient is .70 or .60. However, this is, in most cases, different for descriptive statistics. A small difference in the sample can be enormous on the scale of the whole Dutch population.

## 5 Validity of substantive conclusions generated by the LVO

The LVO has generated several substantive conclusions. We, in respect to the discussion here above, concluded that the descriptive statistics are almost definitely biased and that this bias can have important consequences for the interpretation. With regard to the causal analyses in which *change scores* were employed, we believe that correlations and regression-coefficients are not always unbiased, even when corrections were made. However, we feel that the eventual biases in these estimates present little problems for the interpretation. Moreover, most substantive conclusions were framed in rather general terms (see below), and it is difficult to see how minor changes in the estimates can alter the formulation of these conclusions. The following substantive conclusions which were generated by the Dutch National Mobility Panel and which are known and used in policy-making circles are likely *not* to be affected seriously by the biases discussed above.

1. Travel behaviour tends to change when changes in the life style occur (job promotion, birth of a child, retirement).
2. Changes in mobility are thus related with changes in personal and household conditions (e.g. wage increase, new address). These changes, however, are a-symmetrical. For example, an increase in income leads to a greater increase in mobility, while a decrease in income would not cause as great a decrease in mobility.
3. Effects (of policy measures, a price increase of petrol) on mobility show a certain lag through time. It is important to distinguish between short and long term effects. Longitudinal data are needed to study the development of an effect over time.
4. To some degree travel behaviour (both in terms of preferences for the means of transport and in terms of habit formation in travel behaviour) is stable over time. This is called the stability of travel behaviour.
5. There is no fixed group of transit users. Rather, users of public transport are a 'group' characterized by a high *in-* and *out-flux* of persons who occasionally make use of public transport.

We have the impression that these conclusions are not inflated by the problems and resulting biases that have troubled the analyses of the LVO-data. After all, these conclusions are put in rather general terms and are grounded in causal studies that are unlikely to be affected seriously by the biases.

Some other conclusions or insights are employed, not as concepts or ideas, but in a more instrumental way. We will turn to these conclusions now:

1. The LVO has generated a methodological-technical know-how that have been employed for designing and executing other panel studies (for example the Evaluation of the Flevo Railroad Connection, an impact study of the use of the Zeeburgertunnel and the Transit Users Panel in San Diego, USA).
2. Cross elasticities (the differential effects of, say, a fare increase on public transport use and on car use) were estimated on the LVO-data, and accommodated in MIDAS

(Kitamura & Goulias 1991, draft), a forecasting tool and simulation model, and were compared with similar estimates as used in simulation programs such as the *Landelijk Model Systeem*<sup>10</sup>.

We do not worry about the instrumental use of the first point. The second point, however, is of some concern. The use of biased coefficients of (price and cross) elasticities in simulation models might have some effect on simulated forecasts and for policy purposes this might have certain consequences. In the case of MIDAS, the bias which stemmed from selective attrition is accounted for. Initial non-response and measurement biases due to panel fatigue, however, have not been taken into account, and this might result in biased estimates. On the other hand, the alternative to estimate elasticities from cross sections is not very attractive. Elasticities are better estimated when longitudinal data are used and this is preferred over making longitudinal extrapolations from cross sectional data. Whether small biases do actually have consequences on forecasts depends, of course, on the precise nature of the simulation model and depends on the way simulation results are presented to policy makers.

## 6 A future approach to the problems

As stated before, most of the problems can be seen as a result of the design and data-collection phase of the research. The design itself was constrained by financial- and time budgets and lost focus when goals and expectations kept shifting and changing. Our advice is to limit the number of goals that one wishes to achieve and allow for extra time to develop and test the design (appropriateness of the sample scheme, the method of data collection and the instruments). This will enable consultants to come up with a tailor made design which is more adequate than what we have seen with the LVO. We especially expect to gain much from the following:

1. Restrict the number of topics to be researched in a single study. This indicates that the number of commissioning agencies might have to be limited as well.
2. It is very important to keep the questionnaires as small as possible. This will reduce respondent burden and, subsequently, it will reduce the magnitude of attrition and forms of *panel conditioning* such as systematic underreporting.
3. As a rule, do not use a travel diary. The seven days which were covered in the travel diary used in the LVO are clearly too much of a burden for the respondents. But two days might also be too much as is evident from the Travel Behaviour Study (OVG). Travel diaries can be used only if researchers are absolutely confident that the respondents will conscientiously fill in the travel diaries for the requested period. An example is given by the Telewerk Panel. In the Telewerk Panel a travel diary of seven days was used by a relatively small number of respondents. These respondents were frequently contacted and motivated. Moreover, they were under a contract to give full support to this particular research. This apparently helped.

---

<sup>10</sup> Elasticities were calculated from the LVO-data for the Evaluation of the Fare Increase Plan (cf. Begeleidingsgroep Onderzoek Tarievenplan 1988: 64-67). They were also estimated for MIDAS (see above). In another simulation model, the Landelijk Model Systeem (LMS), estimates of elasticities were based on cross sectional data, and these were used long before comparable estimates became available from the LVO. Agencies working with the LMS-simulation model took notice of the research outcomes of the LVO. It is not certain to what degree estimates of elasticities based on LVO-data have actually been used or are likely to be used in the LMS-model or related simulation models.

4. Reduce respondent burden by using an efficient *routing* in the questionnaires. Also, consider the use of the description of *standard trips* (home - work - home; home - shop - home; home - school - home). Respondents can refer back to these standard trips without the need to describe them over and over again each time.
5. Try to obtain a better understanding of the nature of selective initial non-response by asking refusing respondents to give at least some information on some key topics. This should be done in such a fashion that the non-response can be analyzed and modelled. We recommend that an expert conduct these response - non-response analyses and estimates a non-response model in order to make corrections for selective initial non-response. Such experts are often econometricians.
6. Also include a model estimated for attrition which can be used as a sub-model in the model that one wishes to estimate.
7. The reliability of the measurement of mobility should be improved. Experiments to this effect could be conducted in the Telepanel, using computer-assisted, self administered questionnaires (CASAQ).

#### **7 Some recommendations for longitudinal policy research**

In chapter 4.2 we discussed some of the advantages and disadvantages of several forms of longitudinal research. In consideration of these pro's and con's and with respect to a few common goals of longitudinal policy research in the field of travel behaviour and transport, we have formulated optimal designs for each goal. This resulted in four basic designs (see paragraph 4.2.5 for an extended discussion). Here, we will briefly describe these designs and objectives.

1. Consider a long panel (with, say, 5 or more waves, like the LVO) when the objective is to describe and explain changes on both the aggregate and individual level (*net* and *gross changes*) over time. Consideration should also be given to the points mentioned above in 5.6.
2. Consider a short panel (or 'mini-panel' with a before- and one or two after-measurements) when the objective is to evaluate specific policy interventions, especially when one is also interested in gross changes and when experimental designs can not be implemented.
3. Consider a choice based sample when studying a rare phenomenon (trips by train, car pool behaviour) or when studying a rare sub-population (commuters who travel by public transport). This, however, calls for non-standard statistical procedures.
4. Finally, consider a repeated cross section (like the OVG) when the objective is to describe trends on an aggregate level. The objective then is to monitor trends in general mobility developments.

We like to mention that variations can and should be made on these basic designs according to the specific objectives of the research at hand. Moreover, combinations of the designs can be made. For example, a short panel designed to assess the effectiveness of a locally implemented policy measure can obtain an excellent comparative background when this short panel is combined with a long panel or an ongoing repeated cross sectional survey that monitors trends over a long period. A short panel can be used for evaluating an intervention that is likely to affect only a specific segment of the population and the data for this short panel can be collected from a choice based sample. In any case, experience with the LVO-panel did not indicate that some forms of research should be avoided in the future.