


# Optimal sampling designs for multiday and multiperiod panel survey

**Conference Paper****Author(s):**

Chikaraishi, Yoshito; Fujiwara, Akimasa; Zhang, Junyi; Axhausen, Kay W. ; Zumkeller, Dirk

**Publication date:**

2011

**Permanent link:**

<https://doi.org/10.3929/ethz-a-006909553>

**Rights / license:**

[In Copyright - Non-Commercial Use Permitted](#)

**Originally published in:**

Arbeitsberichte Verkehrs- und Raumplanung 736

## Optimal sampling designs for multiday and multiperiod panel survey<sup>1</sup>

Makoto Chikaraishi (Corresponding author)

Assistant Professor

Graduate School for International Development and Cooperation, Hiroshima University

1-5-1, Kagamiyama, Higashi-Hiroshima 739-8529, Japan

Tel/Fax: +81-82-424-4342; E-mail: [chikaraishi@hiroshima-u.ac.jp](mailto:chikaraishi@hiroshima-u.ac.jp)

Akimasa Fujiwara

Professor

Graduate School for International Development and Cooperation, Hiroshima University

1-5-1, Kagamiyama, Higashi-Hiroshima 739-8529, Japan

Tel/Fax: +81-82-424-6921; E-mail: [afujiw@hiroshima-u.ac.jp](mailto:afujiw@hiroshima-u.ac.jp)

Junyi Zhang

Professor

Graduate School for International Development and Cooperation, Hiroshima University

1-5-1, Kagamiyama, Higashi-Hiroshima 739-8529, Japan

Tel/Fax: +81-82-424-6919; E-mail: [zjy@hiroshima-u.ac.jp](mailto:zjy@hiroshima-u.ac.jp)

Kay W. Axhausen

Professor

Swiss Federal Institute of Technology, Zürich, Switzerland

Tel.: +41-44-633 3943; Fax: +41-44-633 1057

E-mail: [axhausen@ivt.baug.ethz.ch](mailto:axhausen@ivt.baug.ethz.ch)

Dirk Zumkeller

Professor

Institute for Transport Studies, University of Karlsruhe, Germany

Tel. +49-721-6082251; Fax: +49-721-6088031

E-mail: [zumkeller@ifv.uni-karlsruhe.de](mailto:zumkeller@ifv.uni-karlsruhe.de)

**Word Count:** 7,440 words + 9 tables/figures

**Key Words:** multi-day and multi-period panel, sample size, survey duration, frequency of observation, German Mobility Panel, statistical power, non-linear changes

---

<sup>1</sup> Please send any correspondence to Dr. Chikaraishi, Tel./fax: +81-824-24-4342; e-mail: [chikaraishi@hiroshima-u.ac.jp](mailto:chikaraishi@hiroshima-u.ac.jp)

**ABSTRACT**

This paper proposes an optimal survey design method for multiday and multiperiod panels that maximizes the statistical power of the parameter of interest. The method addresses balances among sample size, survey duration for each wave, and frequency of observation. Higher-order polynomial changes in the parameter are also addressed, allowing us to calculate optimal sampling designs for nonlinear changes in response to a given policy intervention. After developing the survey design method and showing numerical simulation results, an empirical analysis is conducted using data from the German Mobility Panel, which is an excellent ongoing multiday and multiperiod survey. In the empirical analysis, we identify optimal survey designs for capturing the impacts of policy interventions on trip generation. One of the most important findings in this study is that variation structure in the behavior of interest strongly influences how surveys are designed to maximize statistical power, while the type of policy to be evaluated does not influence it so much. We also point out several important research issues for the future.

## 1. INTRODUCTION

When a travel (or activity) diary survey is designed, we must determine sampling procedures, questionnaire items, financial schemes, and so on. Academic research in the transportation field has concentrated more on how to use data (e.g., behavioral modeling) than on how to collect it (e.g., survey design). It may be time to pay greater attention to survey design issues, because, unlike model development, data collection is often time sensitive. For some information, if we miss the opportunity to collect the data (such as information that will only be in short-term memory for a limited time), retrospective surveys will not be able to capture the detailed behavioral data of interest. Recent travel surveys have often been smaller in scale (partly because of financial difficulties), and under such situations, we have to carefully consider how to design smaller surveys while minimizing the loss of necessary information. This paper approaches such survey design issues from a statistical perspective.

In the near future, many developed countries will face decreasing populations (indeed, countries such as Germany, Italy and Japan already face the problem), which may trigger a number of microscopic and macroscopic changes. In this situation, it may not be appropriate to apply one-day data to demand forecasting, in which longitudinal trends and changes in behavior are extrapolated from cross-sectional data (Kitamura, 1990). Instead, longitudinal data would be more appropriate for demand forecasting and policy evaluation, especially for cases in which changes are expected over time (see, e.g., Goodwin, 1998). In this context, it is known that multiday and multiperiod panel survey data can provide the information we need to represent behavioral variations and changes in model development and policy evaluation (Pendyala and Pas, 2000). The Dutch Mobility Panel, Puget Sound Transportation Panel, and the German Mobility Panel are prominent examples of multiday and multiperiod panel surveys. These survey data have been used in a number of studies and have not only improved our understanding of activity–travel behavior, but also provided fundamental data that have been used to establish new theoretical foundations. However, there are still relatively few empirical investigations of multiday and multiperiod panels. The reasons for this may include: 1) conducting such longitudinal surveys is seen as more expensive than cross-sectional surveys; 2) such surveys may require more complicated institutional arrangements; and 3) the advantages and disadvantages of applying such complicated survey

data are not clear. Considering these concerns, it would be useful to clarify for a given budget, how, for instance, parameter accuracy is changed by shifting from cross-sectional surveys to multiday and multiperiod surveys, what kind of trade-off structure exists between survey cost reduction and parameter accuracy improvement, and which survey design components (number of sample, survey duration, number of waves, etc.) should be changed to reduce survey cost while minimizing the loss of parameter accuracy. In other words, the development of effective survey design methods could be one way to encourage the use of multiday and multiperiod panel surveys, but this possibility has not been well explored because of limited methodological considerations and limited empirical applications.

We should mention here that several transportation studies have focused on developing effective travel diary survey designs. Pas (1986) established the optimal length (in days) for multiday panel surveys and underscored the substantial benefits of multiday panel surveys for reducing data collection costs and/or improving the precision of parameter estimates. Kitamura et al. (2003) focused on the design of multiperiod panel surveys in the context of discrete travel behaviors, and concluded that continuous behavioral observations are needed to detect changes in behavior. This implies that, to identify changes in behavior, we may have to explicitly distinguish between short-term variability and long-term changes, especially in practical situations (for example, applying one-day data to forecasting involves longitudinal extrapolation of cross-sectional variability). Because multiday data contain information on short-term variability and multiperiod data contain information on long-term changes, using both multiday and multiperiod panel data could be one of the solutions to this problem. However, to our knowledge, there is no empirical research on optimal designs (in terms of survey cost efficiency) for multiday and multiperiod activity–travel diary surveys.

Based on the above considerations, this paper attempts to develop a method for determining optimal design for multiday and multiperiod panel surveys under a given budget constraint, and assuming nonlinear changes in a given parameter. Concretely speaking, we describe the trade-offs between 1) the observed duration of each wave (i.e., how many consecutive days respondents should report their behavior for each wave), 2) the interval between successive waves (i.e., how frequently their behavior is observed), and 3) the number of samples, focusing on the statistical power of the parameter estimate. The proposed method is based on existing methods developed in other fields. In particular, the methodological

framework for managing longitudinal sampling designs developed by Raudenbush and Xiao-Feng (2001) is fundamental to the current study. However, a straightforward application of this method is hampered by the substantial day-to-day variations in travel behavior—such substantial fluctuations of objective variables generally do not appear when this method is used in other fields. To handle such fluctuations, we include a multilevel modeling technique in the Raudenbush and Xiao-Feng (2001) methodological framework, which allows us to distinguish between interindividual and intraindividual variances. This extended method has the same structure as “the cluster randomized trials with repeated measures” described by Spybrook et al. (2011). In this study, we further extend the methodology by introducing a budget constraint.

We believe that this is the first work to focus on the optimal design of multiday and multiperiod travel diary panel surveys. After illustrating the methodological framework for optimal panel survey designs and showing some numerical simulation results, we present an empirical application, using German Mobility Panel data, that focuses on trip generation behavior. Although the findings of this paper are only applicable when the statistical power of a particular parameter is the focus of the survey, the clarified trade-offs among several survey elements could be a useful guide for policy makers who must make difficult survey design decisions.

The next section reviews previous studies focusing on optimal survey designs. In Section 3, a method for optimal survey design of multiday and multiperiod panels is described. Then, following numerical simulations based on the proposed method, we show the empirical results based on the German Mobility Panel. In the final section, key conclusions and future tasks are summarized.

## **2. LITERATURE REVIEW**

### **2.1. Methods for Panel Survey Design**

Although there are some discussions of optimal panel survey design in the transportation field (e.g., Lawton and Pas, 1996; Pendyala and Pas, 2000), there is little methodological research

on optimal survey designs for multiday and multiperiod panels. On the other hand, methodological studies for panel survey designs have been published in other fields, including statistics, psychology, and medical science, as well as in the social, biomedical, and educational research fields. One of the important early studies was done by Hansen *et al.* (1953). They proposed an optimal survey design method for cluster sampling (e.g., cluster randomized trials). Although cluster sampling is known to be inefficient because the data obtained from a given cluster are generally correlated with each other (and thus there is a certain loss of information), Hansen *et al.* showed that this inefficiency may be offset by the reduced survey costs associated with collecting data from the same cluster. In fact, the optimal travel diary survey design method proposed by Pas (1986) is a straightforward extension of Hansen *et al.*'s method to multiday panel surveys, in which the cluster is an individual and each observation is done at the person–day level. Such cluster-sampling-based methods have been further developed in a multilevel modeling approach, especially by researchers in education and psychology (Hox, 2002; Snijders, 2005; Berger and Wong, 2009; Moerbeek *et al.*, 2010). Snijders and Bosker (1993) developed optimal sampling designs for a general two-level linear model, and Raudenbush and Liu (1997) presented optimal survey designs for identifying the effect of a policy invention at the cluster level. Raudenbush and Liu (1997) also showed that a covariate can substantially increase the efficiency of cluster sampling. Moerbeek (2006) introduced a cost function to describe trade-offs between using covariate and increasing sample size. Cohen (1998) developed optimal multilevel survey designs for the estimation of variances of unobserved components, and Cohen (2005) did the same for situations in which intraclass correlation was the primary interest. Moineddin *et al.* (2007) simulated the properties of optimal survey designs for multilevel logistic regression.

The studies reviewed thus far have focused on time-invariant aspects of behavior. Schlesseman (1973) conducted an important initial study of optimal survey designs for changing behavior by using multiperiod panels. The paper examined the proper balance in a longitudinal survey between the frequency of measurements and study duration. The results showed that a unit increase in the study duration reduced the standard error of a parameter estimate more than did a unit increase in the frequency of measurements. Raudenbush and Liu (2001) extended Schlesseman's approach to include higher-order polynomial effects. Bloch (1986) introduced a cost function for optimal multiperiod panel survey designs, and

presented optimal survey designs that tried to strike a balance between additional subjects and additional measurements for each subject. Winkens *et al.* (2005) focused on the optimal time-points for repeated measurements under various covariance structures and found that the commonly used design, with equally spaced measures, is not optimal under certain conditions. Basagaña and Spiegelman (2009) claimed that existing longitudinal research has assumed that exposure (in our context, policy intervention) is time invariant, and they proposed new survey design methods that assumed policy interventions that vary with time. For repeated-measurement survey designs, several studies have addressed panel-specific issues, including dropouts and missing data (e.g., Muthen and Curran, 1997; Galbraith *et al.*, 2002).

## **2.2. Application to Activity–Travel Diary Surveys**

There is also a history of survey designs in the transportation field, for activity–travel diary surveys (Stopher, 2009). In the early period, the sample sizes for home interview surveys were generally defined as a percentage of the population, and often ranged from 1% to 3% of the population [in Japan, slightly higher percentages are used, with calculations based on the idea of Relative Standard Deviation (JSTE, 2008)]. Then, because of increasing survey costs and a better understanding of sampling statistics, sample sizes dropped in the 1970s, and they were no longer calculated as a percentage of the population. Smith (1979) published one of the pioneering works on sample size reduction, claiming that 900–1200 respondents constituted a sufficient sample size. From the 1980s on, multiday panel and/or multiperiod panel surveys have been popular because they allow us to describe dynamic travel behavior with both short- and long-term variability (Pendyala and Pas, 2000). Smart (1984) and Pas (1986) discussed optimal survey designs for multiperiod panels and multiday panels, respectively. However, to the best of our knowledge, there is no transportation research on optimal travel survey designs for multiday and multiperiod panels, although, as mentioned above, there are plenty of panel survey design studies in other fields. One of the crucial reasons why there is little study of this topic in the transportation field might be that there is relatively little data on changes in travel behavior. Such knowledge is generally a prerequisite for determining optimal survey designs (e.g., intuitively speaking, we may need more behavioral observations when there are substantial changes in behavior). In recent years, with



the increasing availability of longitudinal data and the development of modeling methods, a number of studies have explored changes in various aspects of behavior (e.g., Pas, 1987; Pas and Sundar, 1995; Pendyala, 1999; Kitamura et al., 2006; Chikaraishi et al., 2010, 2011). By utilizing these modeling methods, we will conduct empirical studies on optimal survey designs for multiday and multiperiod panels, followed by development of a method for the optimal design of survey panels.

### 3. OPTIMAL SURVEY DESIGN FOR PANEL SURVEYS

The methodological foundation for the current study was formulated by a series of studies in other fields (Schlesseman, 1973; Raudenbush and Liu, 1997; Raudenbush and Xiao-Feng, 2001; Spybrook et al., 2011). Although applying these methods to optimal designs for multiday and multiperiod travel surveys may be worthwhile, we extend this research by introducing a cost function—we set an exact maximization problem under a certain budget constraint.

#### 3.1. Basic Concept

In this paper, the term “survey design” refers to the design of a multiday and multiperiod survey where 1) the observed duration of each wave is denoted by  $D$ , 2) the frequency of survey is denoted by  $F$ , and 3) the number of samples is denoted by  $N$ . Note that the frequency  $F$  is directly related to the overall survey period  $G$  and the total number of waves  $T$ . Concretely,  $F$  can be defined as  $(T-1)/G$ . For example, when the total number of waves  $T$  is 6 and the survey period  $G$  is 10, the frequency  $F$  is equal to  $1/2$ , i.e., the survey would be conducted once every two years. Thus, when two of these terms are set, the remaining element is automatically determined. In this study, for simplicity,  $G$  was a given parameter, and thus, identifying  $F$  is the same as identifying  $T$ . Hereinafter, we mainly use  $T$  instead of  $F$ , but essentially there is no difference.

Based on the above definition of survey design, looking at real examples, in the fourth person trip survey in the Tokyo metropolitan region (which is a kind of traditional person trip

survey), the survey design  $\{N, T, D\} = \{883044, 1, 1\}$  was applied. In the same way, *Mobidrive* survey (Axhausen et al., 2002) was conducted with the survey design  $\{N, T, D\} = \{361, 1, 42\}$ , and the German Mobility Panel (Zumkeller, 2009) was conducted with the survey design  $\{N, T, D\} = \{1800, 17, 7\}$  (in this survey, sample refreshment was applied: each respondent was asked to report a period of continuous one-week travel behavior over each of three years, and thus the sample size fluctuates slightly according to year). The survey designs vary from survey to survey, probably depending on the purpose, and thus determining the purpose is the initial step in survey design. In this study, we set “capturing nonlinear changes in response to a certain policy intervention” as the main purpose of the survey. Specifically, we want to maximize statistical power for the parameter that represents the degree to which policy intervention affects the average rate of change, rate of acceleration, and higher degree polynomial effects. Thus, the term “optimal” survey design refers to a survey design  $\{N, T, D\}$  that maximizes the statistical power of the response to policy intervention. The basic concept for the multiday and multiperiod survey design is presented in Figure 1.

### 3.2. Assumptions

The main assumptions in this study are as follows.

1. The objective variable is continuous.
2. Policy intervention is randomly assigned for  $N/2$  individuals.
3. The survey perfectly follows random sampling procedure.
4. There are no panel-specific problems, such as panel fatigue, dropouts, etc.
5. There are equidistant intervals between successive panels.
6. There is a hierarchical covariance structure (see Subsection 3.3).
7. There is a time-invariant population.

As mentioned in the previous section, a number of methods could be used to relax these assumptions, such as using a logit-type model for Assumption 1 (Moineddin et al., 2007), by introducing an autoregressive covariance structure for Assumption 6 (Winkens et al., 2005),

etc. However, we use these assumptions to make the discussion simple and clear. Future extensions of this method may help to strengthen its practical application to the design of similar surveys.

### 3.3. Model Formula

In this study, the following three-level model is employed:

$$Y_{tdi} = \sum_{p=0}^P \alpha_{pdi} c_{pt} + e_{tdi}, \quad (1)$$

$$\alpha_{pdi} = \beta_{pi} + u_{pdi}, \quad (2)$$

$$\beta_{pi} = \gamma_{p0} + \gamma_{p1} W_i + v_{pi}, \quad (3)$$

where  $Y_{tdi}$  is a dependent variable observed from individual  $i$  ( $= 1, 2, \dots, N$ ), at day  $d$  ( $= 1, 2, \dots, D$ ), in wave  $t$  ( $= 1, 2, \dots, T$ ). Let  $e_{tdi}$ ,  $u_{pdi}$ , and  $v_{pi}$  be normally distributed, with means 0 and variances  $\sigma^2$ ,  $\tau_{\alpha p}$ , and  $\tau_{\beta p}$ , respectively. These random components can be regarded as interwave variation (i.e., variation associated with the repeated measures), intraindividual variation and interindividual variation, respectively. Unknown parameters  $\alpha_{pdi}$ ,  $\beta_{pi}$ ,  $\gamma_{p0}$ , and  $\gamma_{p1}$  have a hierarchical relation:  $\alpha_{pdi}$  is the intraindividual-level coefficient,  $\beta_{pi}$  is the interindividual-level coefficient, and  $\gamma_{p0}$  is the grand mean, for the  $p$ th-order polynomial change parameter  $c_{pt}$ . The term  $\gamma_{p1}$  is the response to policy intervention  $W_i$ , which is a policy intervention indicator set at 1/2 for those who have experienced a policy intervention, and otherwise at  $-1/2$ . Our main interest here is to find optimal survey designs that maximize the statistical power of parameter  $\gamma_{p1}$ . Because the optimal design depends on the order of polynomial change  $p$ , this study derives three different optimal survey designs [i.e., up to the third order of polynomial change ( $P = 3$ )]. To do this, we adopt the method used in Raudenbush and Xiao-Feng (2001) in which orthogonal polynomial contrast coefficients, which allow us to simplify the computations of statistical power, are employed. Specifically, we set orthogonal polynomial contrast coefficients  $c_{pt}$  as follows:

$$c_{0t} = 1, \quad (4)$$

$$c_{1t} = t - \sum_{t=1}^T t / T, \quad (5)$$

$$c_{2t} = \frac{1}{2} \left( c_{1t}^2 - \sum_{t=1}^T t^2 / T \right), \quad (6)$$

$$c_{3t} = \frac{1}{6} \left( c_{1t}^3 - \frac{\sum_{t=1}^T t^4}{\sum_{t=1}^T t^2} c_{1t} \right). \quad (7)$$

### 3.4. Statistical Power

As we mentioned above, in this study, the statistical power of the response to policy intervention  $\gamma_{p1}$  is maximized. We defined the null hypothesis as  $H_0: \gamma_{p1} = 0$ , and the alternative hypothesis as  $H_1: \gamma_{p1} \neq 0$ . The variance of  $\gamma_{p1}$  is defined as follows (see Spybrook et al., 2011):

$$Var(\hat{\gamma}_{p1}) = \frac{4[\tau_{\beta p} + (\tau_{\alpha p} + V_p) / D]}{N}, \quad (8)$$

where

$$V_p = \frac{\sigma^2}{\sum_{t=1}^T c_{pt}^2} = \frac{\sigma^2 F^{2p} (T-p-1)!}{K_p (T+p)!}. \quad (9)$$

$K_p$  is a constant term, where  $K_1 = 1/12$ ,  $K_2 = 1/720$ , and  $K_3 = 1/100,800$ . Here, because statistical power is the probability that the test will reject the null hypothesis when the null hypothesis is false, we can set the probability as follows:

$$P_{1-\beta} = 1 - Pr\{F_0 \geq F(1, N-2; \lambda)\}, \quad (10)$$

$$\lambda = \frac{\gamma_{p1}^2}{Var(\hat{\gamma}_{p1}^2)} = \frac{N\gamma_{p1}^2}{4[\tau_{\beta p} + (\tau_{\alpha p} + V_p) / D]}, \quad (11)$$

where  $F_0$  is the critical value of  $F(1, N-2)$ , which follows the central  $F$  distribution, while  $F(1,$

$N-2; \lambda$ ) follows the noncentral  $F$  distribution with noncentrality parameter  $\lambda$ . Maximizing eq. (10) under a given budget is our main objective here.

### 3.5. Survey Cost Function

Needless to say, larger sample sizes, longer durations, and more frequent measurements will increase statistical power, but they also increase the cost of data collection. Although there are many possible cost functions for multiperiod and multiday panel surveys (depending upon the costs associated with each of these parameters), in this study, we used the following cost function:

$$C = C_0 + C_N N + C_D DN + C_T TN, \quad (12)$$

where  $C$  is a total survey cost,  $C_0$  is the initial setup survey cost,  $C_N$  is the cost for recruiting an individual,  $C_D$  is the cost for increasing an observed duration per individual, and  $C_T$  is the cost for increasing an observed time point (or wave) per individual. Although we have arbitrarily set this cost function, in future research it may be possible to identify survey cost functions through, for example, a kind of meta-analysis.

### 3.6. Maximizing Statistical Power under Budget Constraints

Based on the above-mentioned settings, we set the following maximization problem:

$$\text{Max } P_{1-\beta}(N, T, D) \quad (13)$$

$$\begin{aligned} \text{s.t. } & B \geq C, N > 0, D > 0, T \geq p + 1 \\ & C = C_0 + C_N N + C_D DN + C_T TN \end{aligned} \quad (14)$$

where  $B$  is the total budget that can be used for the survey. The number of waves  $T$  should be bigger than  $p+1$ . This is because when we try to capture a linear change (i.e.,  $p = 1$ ), at least two time-point observations are needed. Thus, this constraint represents the minimum number of waves required to capture the  $p$ -th order polynomial change. The output of this

maximization problem is the optimal survey design  $\{N, T, D\}$ .

We should note here that another optimization problem, i.e., minimizing survey cost while achieving a given level of statistical power, can also be developed in a way similar to the maximization problem described in eqs. (13) and (14). The choice between these optimization methods might depend on the circumstances. In this study, we deal with the situation that a certain fixed survey budget is given.

## 4. NUMERICAL SIMULATION

### 4.1. Basic Settings

Before describing the empirical study, we report numerical simulations we conducted to confirm the behavior of each parameter. The basic parameter settings for the simulations are shown in Table 1. Here, the degree of the response to policy intervention is defined based on the following standardized effect size  $\delta_p$ :

$$\delta_p = \frac{\gamma_{p1}}{\sqrt{\tau_{\beta p} + \tau_{\alpha p}}}. \quad (15)$$

We set the standardized effect size as 0.2, resulting in  $\gamma_{p1} = 0.2 \times (10)^{1/2}$ . In the following subsections, we change one of the parameters in Table 1, and confirm how the parameters affect statistical power. In Subsections 4.2 and 4.3, we calculate statistical power based on eqs. (10) and (11) to confirm the impacts of the changes on the parameters, and in Subsection 4.4 the maximization problem shown in eqs. (13) and (14) is used to check the impacts of changes in cost parameters. Finally, the impacts of effect size on optimal survey designs are identified in Subsection 4.5.

### 4.2. Survey Duration, the Number of Waves, and Sample Size

To begin, the impacts of each survey design element on the statistical power are addressed.

Figure 2 shows the results by the order of polynomial change. From the results, we can confirm first that increases in survey duration, the number of waves and sample size increase statistical power. We also confirm that the higher the order of polynomial change, the longer the survey duration for each wave, the more waves, and/or the larger sample size needed to obtain the same statistical power. This implies that, when a survey is conducted during a period when nonlinear changes can be expected, richer behavioral observation is needed to obtain a given level of statistical power. In addition, the results show that marginal returns for statistical power are basically decreasing with increases in parameters  $N$ ,  $T$ , and  $D$ , whereas a higher-order of polynomial change still keeps higher marginal returns even when the parameters become larger. This indicates that conducting richer multiday and multiperiod surveys would be worthwhile, especially when nonlinear changes are expected.

### 4.3. Unobserved Variations

The impacts of changes in unobserved variations on statistical power are calculated as shown in Figure 3. The results indicate that higher interwave, intraindividual, and interindividual variations consistently reduce statistical power. We also confirmed that, in case of interwave variations, the degree of loss is the greatest with the highest polynomial change (i.e.,  $p = 3$ ). On the other hand, for intra- and interindividual variations, the larger impact is observed in the lower order of polynomial change under the current parameter settings. Based on these results, we can say that, with greater intra- and interindividual variability in behavior, richer multiday and multiperiod survey designs are needed (i.e., longer survey durations for each wave, more waves, and/or larger sample sizes). In addition, because behavioral variability might differ for different aspects of behavior, optimal survey design may strongly depend on the specific aspects of behavior studied, emphasizing the need to clearly define behaviors of interest before designing the survey.

### 4.4. Cost Function

To check the impacts of changes in cost parameters, the parameter settings for the cost function shown in Table 2 were applied as basic settings. How power and the maximization

results vary with cost parameter changes is our interest here.

The maximization results of statistical power and the optimal survey designs under various cost parameters are shown in Figure 4 and Table 3, respectively. From the figure, it can be confirmed that, for the higher-order polynomial change, there are marginal changes in power, and the increases in survey costs are much higher for  $C_D$  and  $C_T$ , compared with  $C_N$ . This implies that, relative to the cost of recruiting individuals, increasing survey durations and waves is more sensitive to the costs under the current cost function. On the other hand, for the 1st-order polynomial change, the sensitivities are not so different among the different types of cost parameters. Thus, when the existence of nonlinear changes is expected, how survey costs are reduced for increasing durations and time points can be crucial, and depending on the cost structure, the optimal survey design could be quite different. This tendency can also be seen in Table 3. For example, when  $C_T$  increases from 100 yen to 1000 yen, the optimal survey design shifts from  $\{N, T, D\} = \{908, 2.0, 4.2\}$  to  $\{473, 2.0, 6.7\}$  for the 1st-order polynomial changes, whereas it shifts from  $\{276, 29.1, 13.9\}$  to  $\{139, 7.9, 25.4\}$  for the 3rd-order polynomial changes. Of course, such discussions are strongly dependent on the cost structure of the survey, which may vary from case to case. Even so, we believe that such theoretical considerations of multiday and multiperiod panel survey data can be a useful guide for those who are designing multiday and multipanel panel surveys.

#### 4.5. Effect Size

The degree of effect size depends on which policy is introduced. For example, regulation-based policies, such as congestion pricing, might have a greater effect size than do nonregulation-based measures, such as information provision. To confirm whether the optimal survey design varies according to the type of policy being evaluated, the impacts of effect size on optimal survey designs are shown in Table 4. From the table, we can confirm that the optimal survey designs are not strongly affected by effect size, implying that the type of policy evaluated through the survey may not be important for survey designs when there is a budget constraint. In other words, the same survey design could be used for evaluating multiple policies.



## 5. EMPIRICAL STUDIES

In this section, we present an empirical example of the proposed survey design method described in Section 3, focusing on the impacts of policy intervention on trip generation. In Subsection 5.1, the empirical data used in this study (from the German Mobility Panel) are briefly described. In Subsections 5.2 and 5.3, the model estimation and optimization results are explained, respectively.

### 5.1. Empirical Data

Data from the German Mobility Panel (Zumkeller, 2009), a multiday and multiperiod panel survey, are used for the empirical analysis. The German Mobility Panel survey has been conducted since 1994. In this survey, each respondent is asked to report a period of continuous one-week travel behavior over each of three years. In our empirical analysis we excluded data from those who dropped out from the survey. We obtained 93,303 samples reported by 4,443 people [i.e., for each respondent, 7 (days) \* 3 (waves) = 21 days of travel behavior were reported] from 1996 to 2008. Because the total survey period  $G$  is 12 years, orthogonal polynomial contrast coefficients can be calculated as  $\{c_{11}, \dots, c_{1t}, \dots, c_{113}\} = \{-6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6\}$ ,  $\{c_{21}, \dots, c_{2t}, \dots, c_{213}\} = \{11, 5.5, 1, -2.5, -5, -6.5, -7, -6.5, -5, -2.5, 1, 5.5, 11\}$ , and  $\{c_{31}, \dots, c_{3t}, \dots, c_{313}\} = \{-11, 0, 6, 8, 7, 4, 0, -4, -7, -8, -6, 0, 11\}$ .

In this empirical analysis,  $Y_{tdi}$  is defined as individual  $i$ 's total number of trips on day  $d$  at wave  $t$  and the policy intervention variable is defined by residential location, i.e., downtown or outskirts. Although residential location itself is not a kind of policy variable, it could be assumed that urban development projects such as TOD (Transit Oriented Development), compact city, etc., influence respondents' mobility levels (represented by trip frequencies) to a greater or lesser extent depending upon where they live. Importantly, the primary purpose for conducting this empirical analysis was to identify the variation structure of trip frequencies. As shown in Subsection 4.5, the degree of the response to this policy variable was not important in the optimal survey design. We could give any value to the parameter for

the policy intervention variable  $\gamma_{p1}$  for the optimal survey designs when the survey budget is fixed, although small fluctuations exist.

## 5.2. Model Estimation Results

The estimation results for the activity generation model are shown in Table 5. They show that location was not a significant factor for all polynomial changes, whereas statistically significant effects of constant variables are observed for the 1st- and 3rd- order polynomial changes, implying that there would be some nonlinear changes in trip frequencies. For the estimation results of random effects, intraindividual and interindividual variations become smaller as the polynomial order increases. To examine the properties of unobserved variations, the following decomposition technique was applied:

$$Var(Y_{idi} | \gamma_{p0}, \gamma_{p1}, W_i) = \underbrace{\sum_{p=0}^P c_{pt}^2 \tau_{\alpha p}}_{\text{intra-individual variation}} + \underbrace{\sum_{p=0}^P c_{pt}^2 \tau_{\beta p}}_{\text{inter-individual variation}} + \underbrace{\sigma^2}_{\text{inter-wave variation}}. \quad (16)$$

The calculated ratio of intraindividual variation, interindividual variation, and interwave variation to the total variation is 11.9%, 29.8%, and 58.3%, respectively. It can be confirmed that, while interindividual variation is higher than intraindividual variation, the biggest unobserved variation is for interwave variation, which is variability associated with repeated measures. This means that the weekly behavior of wave  $t$  differs greatly from that of wave  $t+1$ , implying that a multiperiod survey could be quite important, especially when nonlinear changes are expected (see discussion in Subsection 4.3).

## 5.3. Optimal Survey Designs for Activity Generation

Based on the identified behavioral variations of the trip generation model shown in the previous subsection, here we attempt to derive optimal survey designs with a budget constraint. Optimal survey designs are identified with the following given parameters: Total budget  $B = 5,000,000$  [Japanese yen]; Initial setup survey cost  $C_0 = 2,000,000$  [Japanese yen];

Cost of recruiting an individual  $C_N = 1,000$  [Japanese yen]; Cost of increasing an observed duration per individual  $C_D = 500$  [Japanese yen]; Cost of increasing an observed time point per individual  $C_T = 200$  [Japanese yen]; Total survey period = 12 [years]; and Standardized effect size = 0.2. For the remaining parameters (i.e., parameters for unobserved components), the estimation results shown in the previous section are used. Note that, although the estimated effect size could also be used, we used a given for that parameter because 1) the introduced policy variable (i.e., residential location) was not significant for all polynomial changes, and 2) the effect size has little effect on optimal survey design, as shown in Subsection 4.5.

The optimal survey design for the 1st-, 2nd-, and 3rd-order polynomial changes was identified as  $\{N, T, D\} = \{813, 2.00, 4.58\}$ ,  $\{690, 4.38, 4.94\}$ , and  $\{550, 7.78, 5.80\}$ , respectively, with statistical power = 0.527, 0.421, and 0.373, respectively. As with the numerical simulation results shown in the previous section, for higher-order polynomial changes, not only are more data collection waves needed, but longer multiday periods of behavioral observations per wave are needed as well. Intuitively, this could be understood as a need for distinguishing between short-term and long-term behavioral variability. When complicated nonlinear changes in response to the policy variable can be expected, the behavioral differences between different time-point observations can be explained in two ways: 1) fluctuations/dispersions in behavior, and 2) structural changes in behavioral mechanisms. To distinguish between these two possibilities, more behavioral information in both the near and far term may be needed—that is, making each wave longer (i.e., enrichment of variation information) and increasing the number of waves (i.e., enrichment of change information). On the other hand, a clear distinction between measurement variations and true behavioral changes is very difficult to make, and this may be similar to the ecological fallacy (Robinson, 1950). How close together in time do two observations need to be to have any differences between them regarded as measurement variability rather than actual changes in behavior? Can the temporal averaging of behavior be regarded as typical behavior? This temporal version of the ecological fallacy arises when we attempt to distinguish between variation and change. Although such effects could be minimized when the temporal behavioral rhythms (weekly rhythm, yearly rhythms, etc.) are taken into account, exploring this temporal version of the ecological fallacy may be important in future research, especially

when the existence of substantial nonlinear changes is expected. To do this, existing studies dealing with MAUP (Modifiable Areal Unit Problem), which can be assumed as a geographic version of the ecological fallacy, could be a useful guide (see, e.g., Zhang and Kukadia, 2005).

## 6. CONCLUSION

In designing a multiday and multiperiod survey, there are trade-offs among sample size, survey duration of each wave, and the frequency of observations. In this paper we focused on adjusting these three survey design components to minimize either total survey costs or error in judgment. Specifically, we first developed a survey design method for determining optimal sampling designs for multiday and multiperiod panel surveys with a given budget and in which statistical power for a given parameter was maximized. Nonlinear changes were also taken into account by introducing higher-order polynomial changes. To our knowledge, this is the first study to develop a method for optimal sampling design of multiday and multiperiod travel diary surveys with nonlinear changes in a given parameter. After developing the survey design method and showing numerical simulation results, we conducted an empirical analysis using data from the German Mobility Panel, which is an excellent, ongoing, multiday and multiperiod survey. In our empirical analysis we identified optimal survey designs for capturing the impacts of policy interventions on trip generation.

There are several important findings in the numerical simulations and empirical results. First, we confirmed that when nonlinear changes occur during a survey period, 1) much richer behavioral observation is needed to obtain a given level of statistical power, and 2) survey costs for increasing survey durations and time points can strongly affect optimal survey designs—it could be difficult to decide an optimal survey design without taking into account the data collection cost. Second, in designing a multiday and multiperiod survey, the optimal survey designs may not be affected very much by the effect size, implying that the type of policy to be evaluated would not be important in survey design. Instead, the specific aspect of behavior being surveyed might have bigger impacts on optimal survey design. More precisely, how the behavior of interest varies and changes might be more important for the survey

design with respect to maximizing statistical power. Therefore, because the relationship between data collection and behavioral understanding constrain and influence each other (Axhausen, 2008), deepening our understanding of behavior based on the existing multiday and multiperiod survey data may be an important area for research for the improvement of the transportation planning process. For example, our empirical analysis of trip generation showed that, for interindividual, intraindividual, and interwave variation, the greatest change occurred in interwave variation, implying that weekly behavior at wave  $t$  substantially differs from that at wave  $t+1$ . Such fundamental behavioral understanding could be important for survey design. Finally, this study also highlighted the existence of a temporal version of the ecological fallacy, which may be relatively new in the transportation field. This temporal ecological fallacy may open a new and important future research area, namely, how do we distinguish between behavioral variation and behavioral change. This question may be more important when there are nonlinear changes in behavior, which may make the relationship between variations and changes more complicated.

Of course, this study has a number of limitations. First, we made assumptions in developing our optimal survey designs. Although it may be difficult to eliminate all assumptions, it would be necessary to compare the results obtained with different sets of assumptions to strengthen our findings or to make modifications in our method. For example, addressing panel-specific issues such as attrition bias in the optimal survey design is an important future task to be explored. Second, we only set the maximization problem using a fixed budget, but sometimes a given level of statistical power may be much more important than preserving the budget, such as when the policy intervention is costly, as in road investments. In such a case, minimizing survey costs while maintaining statistical power may be more appropriate. Third, it would be worth identifying appropriate cost functions based on the experiences of existing multiday and multiperiod surveys, perhaps through some form of meta-analysis. Finally, in this study, we identified optimal survey designs for a single parameter. We could extend the proposed method to a multiobjective optimization problem that would incorporate multiple parameters. Such research could strengthen the practical application of the proposed survey design methods.

## Reference

- Axhausen, K. W., Zimmermann, A., Schonfelder, S., Rindsfuser, G., Haupt, T.: Observing the rhythms of daily life: A six-week travel diary, *Transportation*, 29, pp. 95-124, 2002.
- Axhausen, K.W.: Definition of movement and activity for transport modelling, in D.A. Hensher and K.J. Button (eds.) *Handbook of Transport Modelling*, 2nd edition, Elsevier, Oxford, pp. 329-344, 2008.
- Basagaña, X., Spiegelman, D.: Power and sample size calculations for longitudinal studies comparing rates of change with a time-varying exposure, *Statistics in Medicine*, Vol. 29, pp. 181-192, 2009.
- Berger, M.P.F., Wong, W.K.: *An Introduction to Optimal Designs for Social and biomedical Research*, Wiley, 2009.
- Bloch, D.A.: Sample size requirements and the cost of a randomized clinical trial with repeated measurements, *Statistics in Medicine*, Vol. 5, pp. 663-667, 1986.
- Chikaraishi, M., Fujiwara, A., Zhang, J., Axhausen, K. W.: Exploring Variation Properties of Departure Time Choice Behavior Using Multilevel Analysis Approach, *Journal of the Transportation Research Record*, No. 2134, pp. 10-20, 2009.
- Chikaraishi, M., Zhang, J., Fujiwara, A., Axhausen, K. W.: Exploring variation properties of time use behavior based on a multilevel multiple discrete-continuous extreme value model, *Transportation Research Record*, No. 2156, pp. 101-110, 2010.
- Chikaraishi, M., Fujiwara, A., Zhang, J., Axhausen, K.W., Zumkeller, D.: Changes in variations of travel time expenditure: some methodological considerations and empirical results from german mobility panel, *Transportation Research Record*, 2011 (In Press).
- Cohen, M. P.: Determining sample sizes for surveys with data analyzed by hierarchical linear models, *Journal of Official Statistics*, Vol. 14, pp. 267-275, 1998.
- Cohen, M. P.: Sample Size Considerations for Multilevel Surveys, *International Statistical Review*, Vol. 73, pp. 279-287, 2005.
- Galbraith, S., Stat., M., Marschner, I. C.: Guidelines for the design of clinical trials with longitudinal outcomes, *Controlled Clinical Trials*, Vol. 23, pp. 257-273, 2002.
- Goodwin, P.: The end of equilibrium, In T. Gärling and T. Laitila (eds.), *Theoretical Foundations of Travel Choice Modeling*, Elsevier, pp. 103-132, 1998.
- Hansen, M.H., Hurwitz, W.N., Madow, W.G.: *Sample Survey Methods and Theory, Volume II (Theory)*, New York, Wiley, 1953.
- Hox, J.J.: *Multilevel Analysis: Techniques and Applications*, Second Edition, Routledge, New York, 2010.
- JSTE (Japan Society of Traffic Engineers): *Traffic Engineering Handbook*, Japan Society of Traffic Engineers, 2008.
- Kitamura, R.: Panel analysis in transportation planning: An overview, *Transportation Research A*, Vol. 24, pp. 401-415, 1990.
- Kitamura, R., Yamamoto, T., Fujii, S.: The effectiveness of panels in detecting changes in discrete travel behavior, *Transportation Research Part B*, Vol. 37, pp. 191-206, 2003.
- Kitamura, R., Yamamoto, T., Susilo, Y. O., Axhausen, K. W.: How Routine is a Routine? An Analysis of the Day-to-Day Variability in Prism Vertex Location, *Transportation Research Part A*, Vol. 40, pp. 259-279, 2006.
- Lawton, T.K. and Pas, E.I.: Resource paper for survey methodologies workshop, *Proceedings of Conference on Household Travel Surveys: New Concepts and Research Needs*, National Academy Press, Washington D.C., pp. 134-153, 1996.
- Moerbeek, M.: Powe and money in cluster randomized trials: When is it worth measuring a

- covariate?, *Statistics in Medicine*, Vol. 25, pp. 2607-2617, 2006.
- Moerbeek, M., Van Breukelen, G. J. P., Berger, M. P. F.: *Optimal Designs for Multilevel Studies*, In: Leeuw, J., Meijer, E. (eds.), *Handbook of Multilevel Analysis*, Springer, pp. 177-205, 2010.
- Moineddin, R., Matheson, F., Glazier, R. H.: A simulation study of sample size for multilevel logistic regression models, *BMC Medical Research Methodology*, Vol. 7, (<http://www.biomedcentral.com/content/pdf/1471-2288-7-34.pdf>), 2007.
- Muthen, B. O., Curran, P. J.: General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation, *Psychological methods*, Vol. 2, pp. 371-402, 1997.
- Robinson, W.S.: *Ecological Correlations and the Behavior of Individuals*, *American Sociological Review*, Vol. 15, pp. 351-357, 1950.
- Schlesselman, J.: Planning a longitudinal study: II. Frequency of measurement and study duration, *Journal of Chronic disease*, Vol. 26, pp. 561-570, 1973.
- Smart, H.E.: The dynamics of change: Applications of the panel technique to transportation surveys in Tyne and Wear, *Traffic Engineering and Control*, Vol. 25, pp. 595-598, 1984.
- Smith, M.E.: Design of small-sample home-interview travel surveys, *Transportation Research Record*, No. 701, pp. 29-35, 1979.
- Snijders, T.A.B., Bosker, R.J.: Standard errors and sample sizes for two-level research, *Journal of Educational Statistics*, Vol. 18, pp. 237-259, 1993.
- Snijders, T. A. B.: Power and sample size in multilevel modeling, In: Everitt, B. S., Howell, D. C. (eds.), *Encyclopedia of Statistics in Behavioral Science*, Volume 3, Wiley, pp. 1570-1573, 2005.
- Spybrook, J., Raudenbush, S. W., Congdon, R., Martinez, A.: *Optimal Design for Longitudinal and Multilevel Research*, Documentation for the "Optimal Design" Software, 2011.
- Pas, E.I.: Multiday samples, parameter estimation precision, and data collection costs for least squares regression trip-generation models, *Environment and Planning A*, Vol. 18, pp. 73-87, 1986.
- Pas, E. I.: Intrapersonal Variability and Model Goodness-of-Fit, *Transportation Research Part A*, Vol. 21, pp. 431-438, 1987.
- Pas, E. I., Sundar, S.: Intrapersonal Variability in Daily Urban Travel Behavior: Some Additional Evidence, *Transportation*, Vol. 22, pp. 135-150, 1995.
- Pendyala, R.M., Pas, E.I.: Multi-day and multi-period data for travel demand analysis and modeling, *TRB Transportation Research Circular*, E-C008, *Transportation Surveys: Raising the Standard*, 2000.
- Pendyala, R. M.: *Measuring Day-to-Day Variability in Travel Behavior Using GPS Data*, Final Report DTFH61-99-P-00266., FHWA, U.S. Department of Transportation, Washington, D.C. , 1999 (<http://www.fhwa.dot.gov/ohim/gps/index.html>).
- Raudenbush, S.W., Xiao-Feng, L.: Effects of study duration, frequency of observation, and sample size on power in studies of group differences in polynomial change, *Psychological Methods*, Vol. 6, pp. 387-401, 2001.
- Winkens, B., Schouten, H. J. A., van Breukelen, G. J. P., Berger, M. P. F.: Optimal time-points in clinical trials with linearly divergent treatment effects, *Statistics in Medicine*, Vol. 24, pp. 3743-3756, 2005.
- Zhang, M., Kukadia, N.: Metrics of Urban Form and the Modifiable Areal Unit Problem,

Transportation Research Record, No. 1902, pp. 71-79, 2005.  
Zumkeller, D.: The Dynamics of Change – latest results from the German Mobility Panel,  
Paper presented at the 12th International Conference on Travel Behaviour Research,  
Jaipur, India, December 13-18, 2009.



**List of Tables/Figures**

Table 1. Basic Parameter settings for numerical simulation

Table 2. Parameter settings for cost function

Table 3. Optimal survey designs with varying cost parameters

Table 4. Optimal survey designs with varying effect size

Table 5. Estimation results of activity generation model

Figure 1. Basic concept of survey designs for multi-day and multi-period panel

Figure 2. Statistical power vs. survey design components for each order of polynomial change

Figure 3. Statistical power vs. unobserved variations for each order of polynomial change

Figure 4. Maximization results of statistical power under various cost parameters for each order of polynomial change

**Table 1. Basic Parameter settings for numerical simulation**

<b>Survey designs (Objective variables)</b>		
Sample size [person]	$N$	200
Observed duration of each wave [day]	$D$	14
Total number of waves	$T$	6
Total survey period [year]	$G$	5 (i.e., $F=1$ )
<b>Parameters in the model</b>		
Unobserved inter-wave variation	$\sigma^2$	20
Unobserved intra-individual variation	$\tau_\alpha$	5
Unobserved inter-individual variation	$\tau_\beta$	5
Response to policy intervention	$\gamma_{p1}$	$0.2 \times (10)^{1/2}$

**Table 2. Parameter settings for cost function**

<b>Parameters in the cost function [in Japanese yen]</b>		
Total budget	$B$	5,000,000
Initial set-up survey cost	$C_0$	2,000,000
Cost for recruiting an individual	$C_N$	1,000
Cost for increasing an observed duration per individual	$C_D$	500
Cost for increasing an observed time-point per individual	$C_T$	200

**Table 3. Optimal survey designs with varying cost parameters**

Cost parameter			Maximization results (optimal survey designs)								
			$p=1$			$p=2$			$p=3$		
$C_N$	$C_D$	$C_T$	$N$	$D$	$T$	$N$	$D$	$T$	$N$	$D$	$T$
1000	500	100	908	4.2	2.0	576	6.7	8.4	276	13.9	29.1
1000	500	200	815	4.6	2.0	522	7.2	5.8	222	16.1	22.4
1000	500	300	740	4.9	2.0	504	7.2	4.5	196	17.8	18.1
1000	500	400	682	5.2	2.0	468	7.7	4.0	182	19.1	14.8
1000	500	500	628	5.6	2.0	433	8.2	3.7	171	20.5	12.5
1000	500	600	587	5.8	2.0	403	8.6	3.5	164	21.6	10.9
1000	500	700	555	6.0	2.0	379	9.1	3.4	157	22.7	9.7
1000	500	800	526	6.2	2.0	358	9.4	3.3	151	23.6	8.9
1000	500	900	498	6.5	2.0	339	9.8	3.3	145	24.5	8.3
1000	500	1000	473	6.7	2.0	322	10.2	3.2	139	25.4	7.9
1000	100	200	1279	9.5	2.0	964	14.1	3.5	500	31.2	9.4
1000	200	200	1067	7.1	2.0	784	10.2	3.9	364	23.4	12.8
1000	300	200	945	5.9	2.0	677	8.5	4.4	294	19.9	16.2
1000	400	200	869	5.1	2.0	602	7.5	5.0	252	17.5	19.3
1000	500	200	815	4.6	2.0	522	7.2	5.8	222	16.1	22.4
1000	600	200	764	4.2	2.0	489	6.4	6.6	201	14.9	24.9
1000	700	200	731	3.9	2.0	436	6.2	7.8	186	13.6	28.0
1000	800	200	698	3.6	2.0	401	5.9	8.7	174	12.8	29.8
1000	900	200	670	3.4	2.0	375	5.7	9.5	162	12.3	32.1
1000	1000	200	645	3.2	2.0	359	5.3	10.3	154	11.7	34.2
200	500	200	1403	3.1	2.0	828	5.1	4.4	255	14.7	21.1
400	500	200	1171	3.5	2.0	713	5.8	4.6	247	14.9	21.5
600	500	200	1008	3.9	2.0	642	6.1	5.0	237	15.8	20.9
800	500	200	898	4.3	2.0	579	6.6	5.3	229	15.7	22.2
1000	500	200	815	4.6	2.0	522	7.2	5.8	222	16.1	22.4
1200	500	200	741	4.9	2.0	488	7.4	6.2	214	16.8	22.3
1400	500	200	677	5.3	2.0	463	7.6	6.4	210	16.5	23.1
1600	500	200	634	5.5	2.0	426	8.2	6.7	202	17.2	23.1
1800	500	200	593	5.7	2.0	415	8.1	7.0	200	17.2	23.0
2000	500	200	557	6.0	2.0	386	8.7	7.3	192	17.9	23.7

**Table 4. Optimal survey designs with varying effect size**

Effect size	Maximization results (optimal survey designs)											
	$p=1$				$p=2$				$p=3$			
$\delta_p$	$N$	$D$	$T$	Power	$N$	$D$	$T$	Power	$N$	$D$	$T$	Power
0.01	-	-	-	-	-	-	-	-	-	-	-	-
0.02	-	-	-	-	-	-	-	-	-	-	-	-
0.03	-	-	-	-	-	-	-	-	219	15.8	23.9	0.055
0.04	802	4.7	2.0	0.079	534	7.3	4.9	0.070	218	15.6	24.8	0.059
0.05	806	4.6	2.0	0.095	548	7.1	4.7	0.082	222	17.2	19.4	0.064
0.06	802	4.7	2.0	0.116	545	7.1	4.8	0.096	222	17.5	19.0	0.071
0.07	810	4.6	2.0	0.141	545	7.0	4.9	0.113	223	17.0	19.6	0.078
0.08	813	4.6	2.0	0.169	545	7.1	4.9	0.133	221	17.2	20.1	0.087
0.09	813	4.6	2.0	0.202	550	6.9	4.9	0.156	223	17.0	19.9	0.097
0.1	809	4.6	2.0	0.238	545	7.0	5.0	0.182	220	17.0	20.4	0.109
0.11	814	4.6	2.0	0.278	544	7.0	5.2	0.211	225	16.5	20.4	0.122
0.12	815	4.6	2.0	0.321	537	7.1	5.3	0.242	222	16.8	20.6	0.136
0.13	816	4.6	2.0	0.367	536	7.0	5.4	0.276	220	16.9	21.1	0.151
0.14	815	4.6	2.0	0.414	533	7.1	5.4	0.312	225	16.2	21.2	0.168
0.15	815	4.6	2.0	0.463	538	7.0	5.4	0.350	225	16.2	21.3	0.186
0.16	815	4.6	2.0	0.513	528	7.0	6.0	0.391	224	16.0	22.0	0.206
0.17	816	4.5	2.0	0.562	533	7.0	5.7	0.431	225	15.9	22.0	0.226
0.18	816	4.6	2.0	0.611	535	6.9	5.9	0.474	225	16.0	21.8	0.248
0.19	814	4.6	2.0	0.658	537	6.9	5.6	0.515	220	16.2	22.6	0.271
0.2	815	4.6	2.0	0.702	522	7.2	5.8	0.556	222	16.1	22.4	0.295

Note) “-” means the optimal solution cannot be identified because the power is too small.

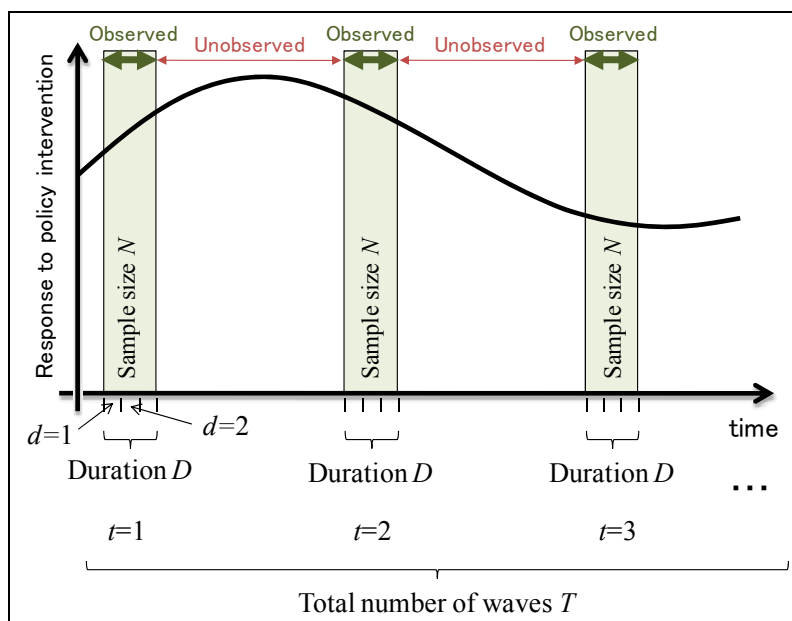
**Table 5. Estimation results of activity generation model**

Item		parameter	t-value
<i>Explanatory variables</i>			
$c_{0t}$	constant	$\gamma_{00}$	3.543
	living in inner city [D]	$\gamma_{01}$	0.008
$c_{1t}$	constant	$\gamma_{10}$	-0.015
	living in inner city [D]	$\gamma_{11}$	0.000 <sup>1)</sup>
$c_{2t}$	constant	$\gamma_{20}$	0.005
	living in inner city [D]	$\gamma_{21}$	-0.003
$c_{3t}$	constant	$\gamma_{30}$	-0.008
	living in inner city [D]	$\gamma_{31}$	0.001
<i>Random effects</i>			
$c_{0t}$	intra-individual variation	$\tau_\alpha$	0.548
	inter-individual variation	$\tau_\beta$	0.927
$c_{1t}$	intra-individual variation	$\tau_\alpha$	0.006
	inter-individual variation	$\tau_\beta$	0.019
$c_{2t}$	intra-individual variation	$\tau_\alpha$	0.000 <sup>2)</sup>
	inter-individual variation	$\tau_\beta$	0.006
$c_{3t}$	intra-individual variation	$\tau_\alpha$	0.000 <sup>3)</sup>
	inter-individual variation	$\tau_\beta$	0.004
	inter-wave variation	$\sigma^2$	3.120
Initial log likelihood (only constant)			-208,487
Final log likelihood			-198,033
Number of sample			93,303

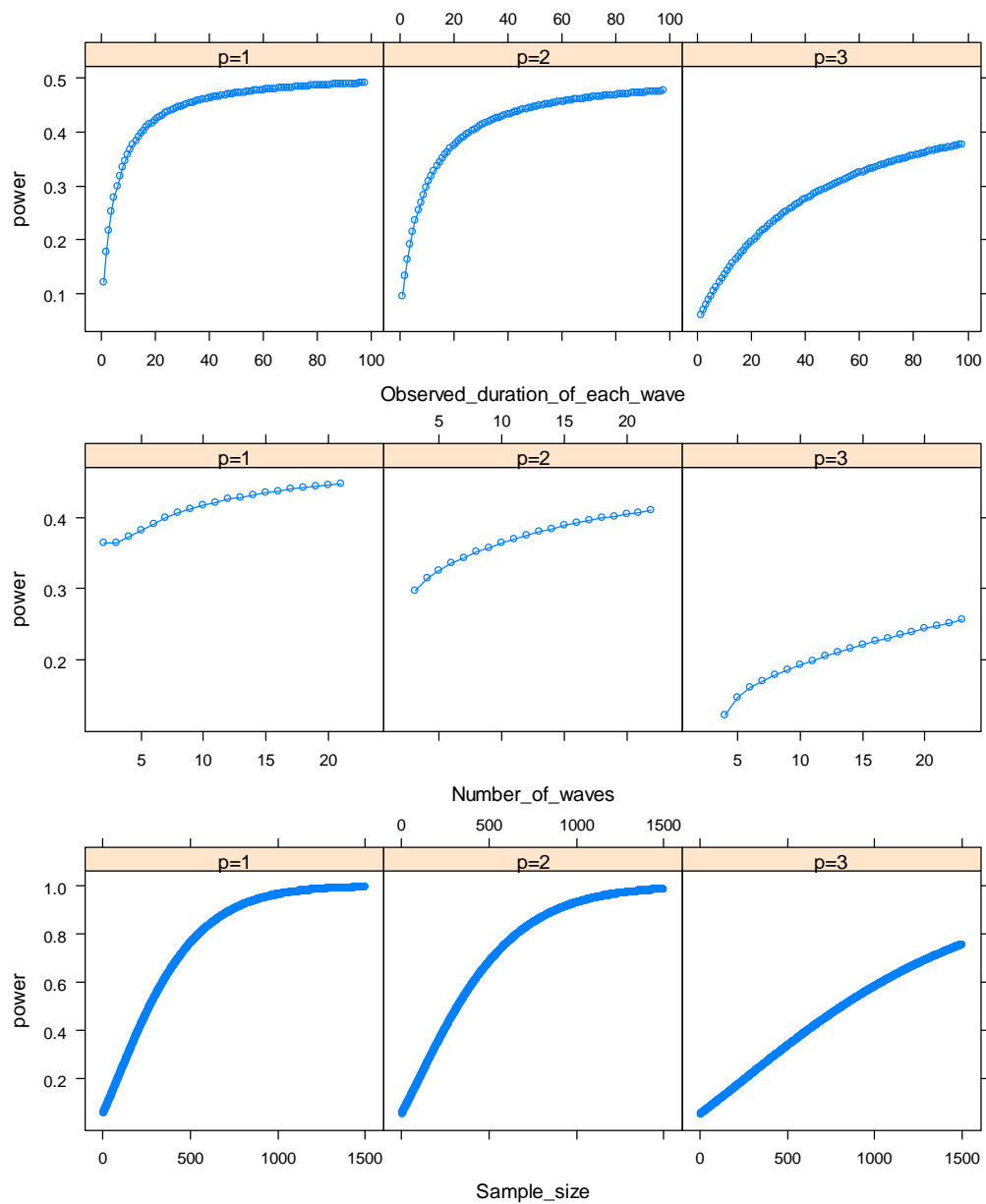
1) The estimated value is -0.000448.

2) The estimated value is 0.0000000000846.

3) The estimated value is rounded to zero in the model estimation.

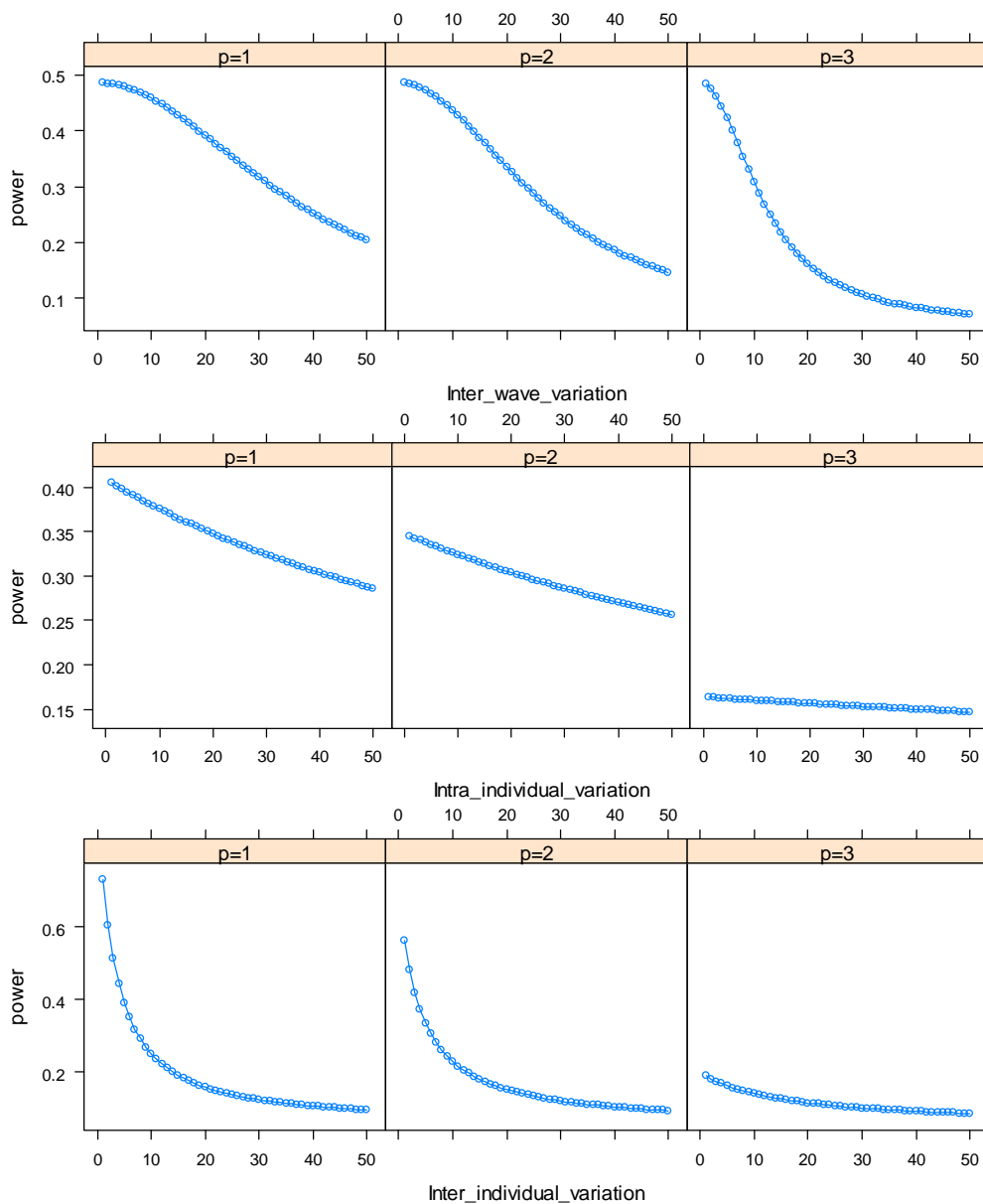


**Figure 1.** Basic concept of survey designs for multi-day and multi-period panel

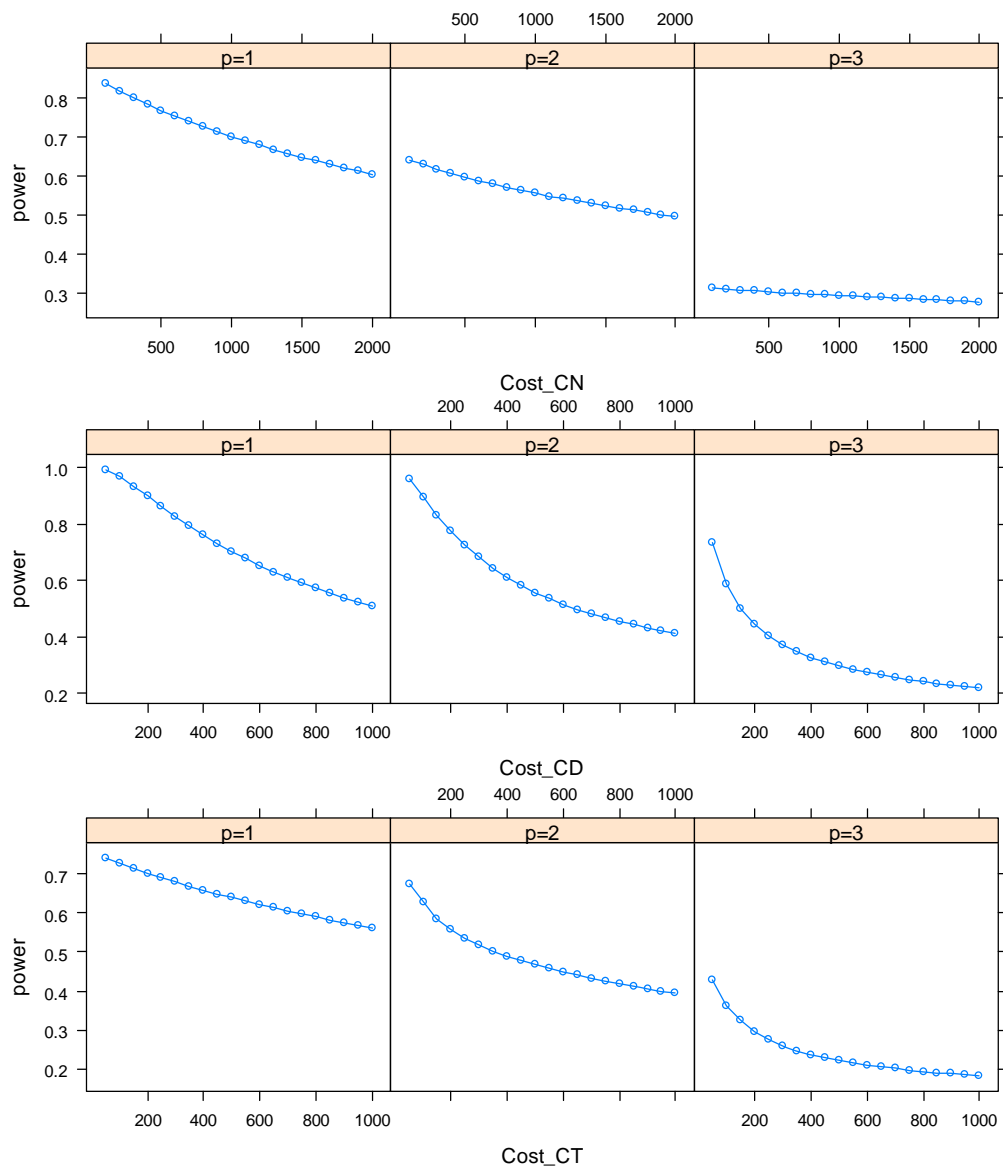


**Figure 2. Statistical power vs. survey design components for each order of polynomial change**





**Figure 3. Statistical power vs. unobserved variations for each order of polynomial change**



**Figure 4. Maximization results of statistical power under various cost parameters for each order of polynomial change**