A time series approach to estimate discontinuities due to a survey redesign

09

Jan van den Brakel and Joeri Roels

Discussion paper (09003)

Statistics Netherlands

The Hague/Heerlen, 2009

Explanation of symbols

.	= data not available
*	= provisional figure
x	= publication prohibited (confidential figure)
–	= nil or less than half of unit concerned
–	= (between two figures) inclusive
0 (0,0)	= less than half of unit concerned
blank	= not applicable
2005-2006	= 2005 to 2006 inclusive
2005/2006	= average of 2005 up to and including 2006
2005/'06	= crop year, financial year, school year etc. beginning in 2005 and ending in 2006
2003/'04–2005/'06	= crop year, financial year, etc. 2003/'04 to 2005/'06 inclusive

Due to rounding, some totals may not correspond with the sum of the separate figures.

Publisher
Statistics Netherlands
Henri Faasdreef 312
2492 JP The Hague

Prepress
Statistics Netherlands - Facility Services

Cover
TelDesign, Rotterdam

Information
Telephone .. +31 88 570 70 70
Telefax .. +31 70 337 59 94
Via contact form: www.cbs.nl/information

Where to order
E-mail: verkoop@cbs.nl
Telefax .. +31 45 570 62 68

Internet
www.cbs.nl

ISSN: 1572-0314

© Statistics Netherlands, The Hague/Heerlen, 2009.
Reproduction is permitted. 'Statistics Netherlands' must be quoted as source.

6008309003 X-10

A time series approach to estimate discontinuities due to a survey redesign

Jan van den Brakel and Joeri Roels

Summary: An important quality aspect of official statistics produced by national statistical institutes is comparability over time. To maintain uninterrupted time series, surveys conducted by national statistical institutes are often kept unchanged as long as possible. To improve the quality or efficiency of a survey process, however, it remains inevitable to adjust methods or redesign this process from time to time. Adjustments in the survey process generally affect survey characteristics such as response bias and therefore have a systematic effect on the parameter estimates of a sample survey. Therefore it is important that the effects of a survey redesign on the estimated series are explained and quantified. In this paper a structural time series model is applied to estimate discontinuities in series of the Dutch survey on social participation and environmental consciousness due to a redesign of the underlying survey process.

Keywords:intervention analysis, response bias, structural time series models, survey sampling

1. Introduction

In 1997 Statistics Netherlands started the Permanent Survey on Living Conditions (PSLC). This is a module-based integrated survey combining various themes concerning living conditions and quality of life. Two modules of the PSLC are called the Module Justice and Environment and the Module Justice and Participation. Both modules are used to publish figures about justice and crime victimisation. The first module is also used to publish figures about environmental consciousness. The second module is used additionally to publish information about social participation. To realize expenditure cuts, the PSLC stopped at the end of 2004. From that moment on, figures about social participation and environmental consciousness are based on a separate survey, called the Dutch Survey on Social Participation and Environmental Consciousness (SSPEC).

In this survey transition the data collection mode, the questionnaire, the context of the survey and the fieldwork period changed, which resulted in systematic effects in the outcomes of the survey. Quality of official statistics is based on various dimensions; see Brackstone (1999) for a discussion. One important quality aspect is comparability over time. To produce consistent series, national statistical institutes generally keep their survey processes unchanged as long as possible. It remains inevitable, however, to redesign survey processes from time to time. In an ideal survey transition process, the systematic effects of the redesign are explained and quantified in order to keep series consistent and preserve comparability of the outcomes over time.

There are various possibilities to quantify the effect of a survey redesign, see Van den Brakel, Smith and Compton (2008) for an overview. In this application the redesign mainly affects the data collection process. In such cases a large scale field experiment is very appropriate to test the effect of the redesign on the parameter estimates of the survey, see e.g. Van den Brakel (2008). An experimental approach might, however, be hampered due to budget and other practical constraints, which was the case for the Dutch SSPEC. In such cases an intervention analysis using a structural time series model can be used as an alternative to quantify the effect of the redesign on the main series of the sample survey.

In section 2 the PSLC and the SSPEC are described. The systematic effects due to the redesign are discussed in section 3. A time series model to quantify these discontinuities is developed in section 4. Results for the most important indicators are given in section 5. The paper concludes with a discussion in section 6.

2. Survey designs

2.1 Permanent Survey on Living Conditions

The PSLC is a module-based integrated survey combining various themes concerning living conditions and quality of life. This survey has been conducted by Statistics Netherlands from 1997 until 2004. The Module Justice and Environment is used to collect statistical information about justice and crime victimisation and environmental consciousness. The Module Justice and Participation is used to collect statistical information about justice and crime victimisation and social participation. Both modules use persons aged 15 years or older as the target population. The sample design is based on stratified two-stage sampling of persons and is self-weighted. The PSLC was a continuously conducted survey with a yearly net sample size of about 4000 to 5000 persons for both modules. The response rates varied around a level of 60%. Interviewers visited all the sampled persons at home and administered the questionnaire in a face-to-face interview, generally referred to as computer assisted personal interviewing (CAPI). The estimation procedure used to compile official statistics is based on the generalised regression estimator (Särndal et al. 1992, chapter 6) using a weighting scheme that is based on different sociodemographic categorical variables.

2.2 Survey on Social Participation and Environmental Consciousness

The PSLC stopped at the end of 2004. From that moment figures about social participation and environmental consciousness are based on the SSPEC. This survey is based on a self-weighted stratified two-stage sample design of persons aged 15 years and older residing in the Netherlands. Data are collected by computer assisted telephone interviewing (CATI). As a result the target population changed from persons aged 15 years or older to the population aged 15 years and older with a non-secret landline telephone connection or cell-phone number. The data collection of the SSPEC is conducted in the months September, October and November. The estimation procedure is, like the PSLC, based on the generalised regression estimator. The response rates in the SSPEC varied around 65%. About 4000 to 4500 respondents are observed in the yearly samples.

Since 2005, figures about justice and crime victimisation are based on the Dutch Security Monitor. See Van den Brakel, Smith and Compton (2008) for more details about this redesign and the effects on the main series of this survey.

2.3 Target parameters

All target parameters about environmental consciousness and social participation are based on closed questions where the respondent can choose one out of K answer categories to specify his opinion or behaviour on an ordinal scale. The target parameters are the estimated proportions that specify the distribution over these K categories for the entire population or subpopulations.

Typical parameters for environmental behaviour are:

- Bringing glass waste to the bottle bank, separating garden waste, vegetable and fruit waste, chemical waste, paper waste, with five answer categories: 1) always, 2) often, 3) sometimes, 4) rarely, 5) never.
- Suffer from noise pollution from airplanes, trains, motorised traffic, factories, loading and unloading cargo, neighbours, suffer from smell pollution from motorised traffic, factories, farms, fireplaces and multi-burners in three categories: 1) yes, 2) sometimes, 3) no.

- Frequency of using a bike instead of a car for travelling short distances in five categories: 1) always, 2) often, 3) sometimes, 4) hardly, 5) never.

Typical parameters for social participation are:

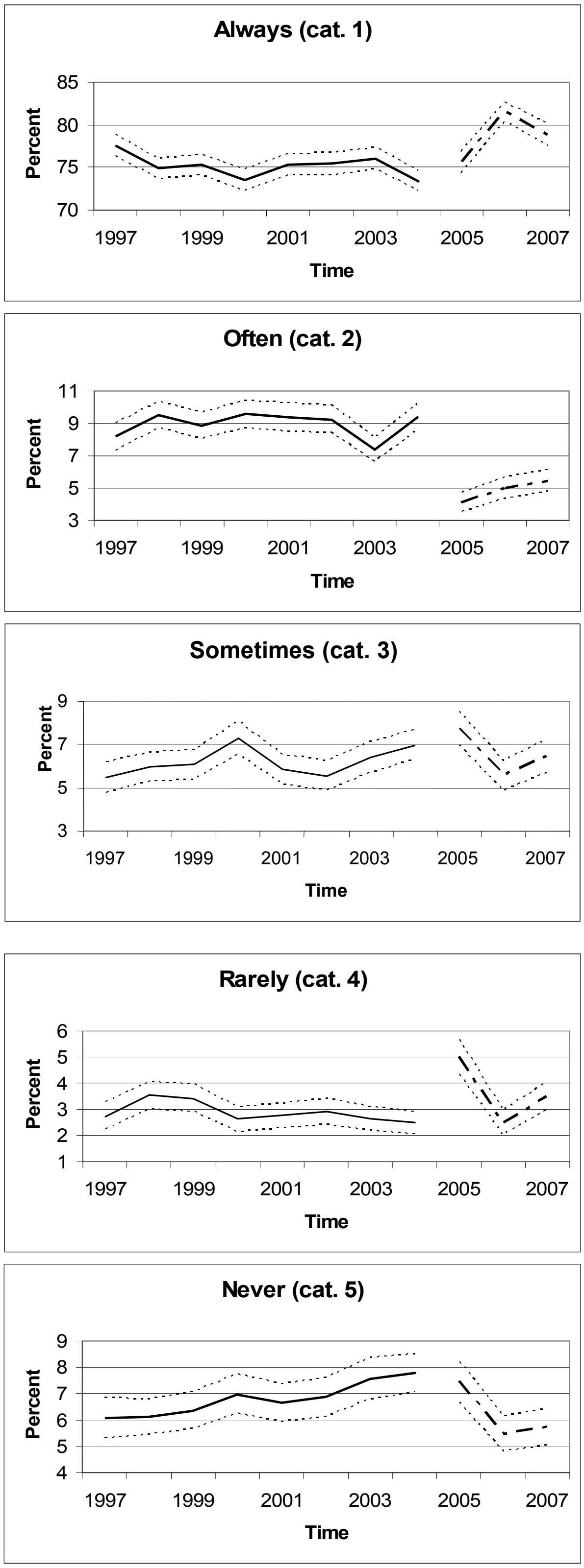
- Time per week spent on mind games and do-it-yourself jobs in four categories: 1) more than 5 hours, 2) 1-5 hours, 3) less than 1 hour, 4) never.
- Frequency of visiting a dance evening, cinema, theatre, concerts and opera, making trips in the countryside, playgrounds, restaurants in four categories: 1) at least once a month, 2) three times a year, 3) less than three times a year, 4) never
- Participation in activities for school, sports club, labour union, hobby club, culture club, youth work, welfare or nursery work in two categories: 1) yes, 2) no.
- Member of a library, sports club, hobby club, labour union in two categories: 1) yes, 2) no.
- Frequency of contact with neighbours in four categories: 1) at least once a week, 2) once within two weeks, 3) less than once within two weeks, 4) never.
- Opinion about feeling socially isolated, member of circle of friends, speaking with other people in three categories: 1) yes, 2) sometimes, 3) no.
- The question "Are your social contacts superficial?", and the statement "There are people who have sympathy for my feelings and opinions" in three categories; 1) yes, 2) sometimes, 3) no.
- Satisfaction about leisure activities, circle of friends and acquaintances in five categories: 1) exceptionally satisfied, 2) very satisfied, 3) satisfied, 4) rather satisfied, 5) not satisfied.

3. Factors responsible for discontinuities

The redesign from the PSLC to the SSPEC resulted in discontinuities in most of the parameters about social participation and environmental consciousness. As an example the series with the annual figures of the parameters separating chemical waste and contact frequency with neighbours are shown in Figures 1 and 2 respectively. For both parameters it appears that there are significant discontinuities in two or more of the underlying categories. The observed differences during the transition from the PSLC in 2004 to the SSPEC in 2005 for all other parameters about environmental consciousness and social participation are summarised in

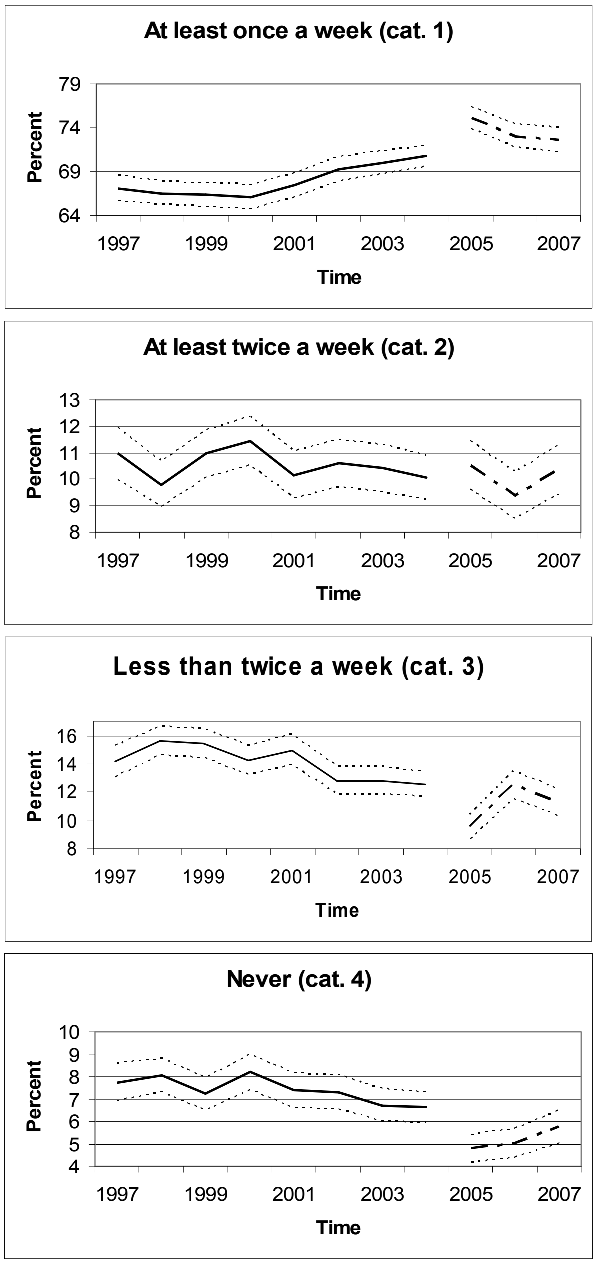
Annex A. The standard errors are approximated with a variance estimator for fractions; see e.g. Cochran (1977), Ch. 3.

Figure 1: Separating chemical waste



Solid line: observed series under the PSLC, dashed line: observed series under the SSPEC, dotted line: 95% confidence interval.

Figure 2: Contact frequency with neighbours



Solid line: observed series under the PSLC, dashed line: observed series under the SSPEC, dotted line: 95% confidence interval.

The observed differences are the results of the factors that changed simultaneously in the survey redesign, real developments of the parameter and sampling errors. The most important factors that changed in the survey redesign are:

- Differences between sampled target populations. The SSPEC is based on a sample of persons aged 15 years and older with a non-secret permanent telephone connection or cell-phone number. The PSLC is based on a sample of all persons aged 15 years and older. The SSPEC does not observe the subpopulation that does not have a non-secret telephone number. Additional analyses showed that this results in an under-representation of young people and ethnic minorities. This explains a substantial part of the discontinuities.

- Differences in data collection modes. The SSPEC is a telephone based survey, while in the PSLC data are collected in face-to-face interviews conducted at the respondents' homes. Many references in the literature emphasize that different collection modes have systematic effects on the responses, see for example De Leeuw (2005) and Dillman and Christian (2005). Systematic differences in the data due to the fact that the interviews are conducted by telephone instead of face-to-face are called mode effects. They arise for different reasons. Generally the interview speed in a face-to-face interview is lower compared to an interview conducted by telephone. Further more, respondents are more engaged with the interview and are more likely to exert the required cognitive effort to answer questions carefully in a face-to-face interview. Also fewer socially desirable answers are obtained under the CAPI mode due to the personal contact with the interviewer. As a result, fewer measurement errors are expected under the CAPI mode (Holbrook e.a., 2003, and Roberts, 2007).
- Differences between data collection periods. The data collection for the SSPEC is conducted in September through November, while the PSLC is conducted continuously throughout the year. In the series of the quarterly figures observed under the PSLC, seasonal effects are observed in several parameters, which partially explain the discontinuities.
- Differences between questionnaire designs. Under the PSLC, questions about social participation and environmental consciousness were combined with questions about justice and crime victimisation in two different modules. Under the SSPEC, the questions about social participation and environmental consciousness are delineated in a new survey, which might have systematic effects on the outcomes of these surveys (Kalton and Schuman, 1982 and Dillman and Christian, 2005).
- Differences between the contexts of the surveys. The SSPEC is introduced as a survey that is focussed on topics about social participation and environmental consciousness. The PSLC is introduced as a more general survey on living conditions. Subsequently the survey focuses on topics about justice, crime victimisation, social participation or environmental consciousness. This might have a systematic selection effect on the respondents who decide to participate in the survey. Furthermore, in the SSPEC the attention of the respondent is completely focussed on one topic,

contrary to the PSLC, which also may have systematic effects on the answer patterns of the respondents.

It is not immediately clear to what extend the differences summarized in Annex A are the result of a real development or are induced by the redesign of the survey. Even if no significant difference is observed, it is still possible that a real development could be nullified by an opposite redesign effect.

A general way to avoid confounding the autonomous development with redesign effects is to conduct an experiment embedded in the ongoing survey, where the regular and new approaches are run concurrently for some period. Depending on the available resources, a two treatment experiment can be conducted to separate the net effect of all factors that changed simultaneously form the real development of the parameter under consideration. If the effect of the separate factors that has varied in the survey process should be quantified, then a factorial design should be considered. See Van den Brakel (2008) and Van den Brakel, Smith and Compton (2008) for a detailed discussion and alternative approaches to quantify the effect of a survey redesign.

Since an experimental approach is not applied in this application, a time series model is developed in this paper to separate the real development from the net effect of the survey redesign.

The time series modelling approach as well as a two treatment experiment are appropriate to quantify the net effect of the survey redesign with the purpose to avoid confounding with real developments of the respective parameter. Explaining these observed differences generally requires insight in the effect of the separate factors that changed in the survey redesign. This can be obtained by running a factorial experiment, from subject matter knowledge and findings in the literature. Alternative calculations can also be considered to obtain some insight in the effect for some of the factors that have changed in the survey redesign. The selection effect of surveying the subpopulation that can be contacted by telephone can be estimated from the PSLC since this survey approaches the entire population face-to-face. Telephone numbers, however, are not standard available for the respondents in the samples of the PSLC, so this requires searching telephone numbers for respondents observed in samples from the past, which is very expensive. The effect of changing the period of data collection can also be quantified by making, for example, quarterly series for the PSLC and estimate the seasonal pattern. Due to the relatively small sample sizes and the limited length of the series, it turned out to be hard to establish significant seasonal effects.

4. Time series models

In this section structural time series models are developed to estimate the discontinuities in the series of a survey due to the redesign of the underlying survey process. With a structural time series model, a series is decomposed in a trend component, seasonal component, other cyclic components, regression component and an irregular component. For each component a stochastic model is assumed. This allows the trend, seasonal, and cyclic component but also the regression coefficients to be time dependent. If necessary ARMA components can be added to capture the autocorrelation in the series beyond these structural components. See Harvey (1989) or Durbin and Koopman (2001) for details about structural time series modelling.

4.1 Intervention analysis

In the case of the PSLC and the SSPEC a relatively short series for annual data is considered. Therefore, the autonomous development of the indicator that is described by the series is modelled with a stochastic trend, a regression component and an irregular component. The regression component consists of an intervention variable with a time independent regression coefficient that describes the effect of the survey transition. Seasonal, cyclic, ARMA, and other auxiliary regression components can be included in the model for example in the case of longer series or monthly or quarterly data.

For each parameter *K* series are observed that specify the estimated proportions over *K* categories and are collected in the *K*-dimensional vector $\hat{\mathbf{y}}_t = (\hat{y}_{t,1}, \ldots \hat{y}_{t,K})$, $t = 1, \ldots, T$. The univariate structural time series model for the *k*-th component of $\hat{\mathbf{y}}_t$ is defined as:

$$\hat{y}_{t,k} = L_{t,k} + \beta_k \delta_t + \varepsilon_{t,k}\,, \tag{1}$$

with $L_{t,k}$ a stochastic trend, δ_t an intervention variable that describes under which survey the observations are obtained at period *t*, β_k the time independent regression coefficient for the intervention variable and $\varepsilon_{t,k}$ an irregular component. The irregular component is assumed to be normally and independently distributed with zero mean and a variance that is proportional to the sample size:

$$\varepsilon_{t,k} \cong N(0, \frac{\sigma^2_{\varepsilon,k}}{n_t})\,.$$

It is also assumed that the irregular components of (1) at different time points are uncorrelated: $Cov(\varepsilon_{t,k}\varepsilon_{t',k}) = 0$ for $t \neq t'$. For the stochastic trend, the widely applied smooth trend model is assumed, see e.g. Durbin and Koopman, (2001):

$$L_{t,k} = L_{t-1,k} + R_{t-1,k} \ ,$$
$$R_{t,k} = R_{t-1,k} + \eta_{t,R,k} \ , \qquad (2)$$

with $L_{t,k}$ the level component and $R_{t,k}$ the stochastic slope component of the trend, and $\eta_{t,R,k}$ irregular components. It is assumed that the irregular components are normally and independently distributed:

$$\eta_{t,R,k} \cong N(0, \sigma^2_{R,k}) \ .$$

The irregular components of (2) at different time points are assumed to be uncorrelated: $Cov(\eta_{t,k}\eta_{t',k}) = 0$ for $t \neq t'$. Furthermore, it is assumed that the irregular components of (1) and (2) are uncorrelated: $Cov(\varepsilon_{t,k}\eta_{t',k}) = 0$ for all t and t'. The intervention variable models the effect of the survey redesign. Two types of interventions are discussed: a level shift and a slope intervention. Let T_R denote the time period at which the survey process is redesigned. In the case of a level intervention it is assumed that the magnitude of the discontinuity due to the survey redesign is constant over time. In this case δ_t is defined as a dummy variable:

$$\delta_t = \begin{cases} 0 & \text{if} \quad t < T_R \\ 1 & \text{if} \quad t \geq T_R \end{cases} \qquad (3)$$

In the case of a slope intervention it is assumed that the magnitude of the discontinuity increases over time. This is accomplished by defining δ_t as:

$$\delta_t = \begin{cases} 0 & \text{if} \quad t < T_R \\ 1 + t - T_R & \text{if} \quad t \geq T_R \end{cases} \qquad (4)$$

Other types of interventions are also possible, for example an intervention on the seasonal or cyclic pattern. Under the assumption that the stochastic trend model approximates the real development of the parameter reasonably well, the estimate for the regression coefficient of the intervention variable can be interpreted as the discontinuity in the series due to the survey redesign.

The general way to proceed is to put this model in state space representation and analyse the model with the Kalman filter. The state space representation for this *K*-dimensional structural time series model reads as:

$$\hat{\mathbf{y}}_t = \mathbf{Z}_t \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t \qquad (5)$$

$$\boldsymbol{\alpha}_t = \mathbf{T}\boldsymbol{\alpha}_{t-1} + \boldsymbol{\eta}_t \qquad (6)$$

The measurement equation (5) describes how the observed series depends on a vector of unobserved state variables $\boldsymbol{\alpha}_t$ and a vector with disturbances $\boldsymbol{\varepsilon}_t$. The state vector contains the level and slope components of the trend models and the regression coefficients of the intervention variables. The transition equation (6) describes how these state variables evolve over time. The vector $\boldsymbol{\eta}_t$ contains the disturbances of the assumed first-order Markov processes of the state variables. The matrices in (5) and (6) are given by

$$\boldsymbol{\alpha}_t = (L_{t,1}, R_{t,1}, ..., L_{t,K}, R_{t,K}, \beta_1, ..., \beta_K)^T, \tag{7a}$$

$$\mathbf{Z}_t = \left(\mathbf{I}_{[K]} \otimes (1,0) \mid \delta_t \mathbf{I}_{[K]}\right), \tag{7b}$$

$$\mathbf{T} = Blockdiag(\mathbf{T}_{tr}, \mathbf{I}_{[K]}), \tag{7c}$$

$$\mathbf{T}_{tr} = \mathbf{I}_{[K]} \otimes \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \tag{7d}$$

with $\mathbf{0}_{[p]}$ a column vector of order p with each element equal to zero, $\mathbf{1}_{[p]}$ a column vector of order p with each element equal to one, and $\mathbf{I}_{[p]}$ the $p \times p$ identity matrix. The disturbances vectors are defined as

$$\boldsymbol{\varepsilon}_t = (\varepsilon_{t,1}, ..., \varepsilon_{t,K})^T,$$

$$\boldsymbol{\eta}_t = (0, \eta_{t,R,1}, ..., 0, \eta_{t,R,K}, \mathbf{0}_{[K]}^T)^T.$$

It is assumed that

$$E(\boldsymbol{\varepsilon}_t) = \mathbf{0}_{[K]},\ Cov(\boldsymbol{\varepsilon}_t) = \frac{1}{n_t} Diag(\sigma_{\varepsilon 1}^2, ..., \sigma_{\varepsilon K}^2),$$

$$E(\boldsymbol{\eta}_t) = \mathbf{0}_{[3K]},\ Cov(\boldsymbol{\eta}_t) = Diag(0, \sigma_{R1}^2, ..., 0, \sigma_{RK}^2, \mathbf{0}_{[K]}^T).$$

In the case that each measurement equation and each transition equation has its own separate hyper parameter, then (7) is a set of K univariate structural time series models. If the measurement equations or the transition equations share common hyper parameters, then (7) is a K dimensional seemingly unrelated multivariate structural time series model. This is for example the case if $\sigma_{\varepsilon 1}^2 = ... = \sigma_{\varepsilon K}^2 = \sigma_{\varepsilon}^2$. After having expressed the multivariate structural time series model in state space representation, the Kalman filter is applied to obtain optimal estimates for the state variables as well as the measurement equation see e.g. Durbin and Koopman, (2001).

In this application, each parameter specifies the proportions over K categories. In other words, each parameter makes up a K-dimensional series, which obeys the restriction that at each point in time these series add up to one, i.e. $\sum_{k=1}^{K} \hat{y}_{t,k} = 1$, and $0 \le \hat{y}_{t,k} \le 1$. As a result, the K regression coefficients of the intervention variables must obey the restriction $\sum_{k=1}^{K} \beta_k = 0$. The multivariate structural time series model (7) can be augmented with this restriction by using the following design matrix in the transition equation (6):

$$\mathbf{T} = Blockdiag(\mathbf{T}_{tr}, \mathbf{T}_{iv}), \tag{7e}$$

where $\mathbf{T}_{tr}$ is defined by (7d), and

$$\mathbf{T}_{iv} = \begin{pmatrix} \mathbf{I}_{[K-1]} & \mathbf{0}_{[K-1]} \\ -\mathbf{1}_{[K-1]}^{T} & 0 \end{pmatrix}. \tag{7f}$$

Due to $\mathbf{T}_{iv}$, defined in (7f), the regression coefficients as well as their Kalman filter estimates obey the restriction $\sum_{k=1}^{K} \beta_k = 0$. In the case of a level intervention, the time series after the moment of the survey transition can be adjusted for the estimated discontinuities with $\tilde{y}_{t,k} = \hat{y}_{t,k} - \hat{\beta}_k$. As an alternative, the series before the survey transition can be adjusted with $\tilde{y}_{t,k} = \hat{y}_{t,k} + \hat{\beta}_k$. In the case of a slope intervention the time series is adjusted with $\tilde{y}_{t,k} = \hat{y}_{t,k} - \hat{\beta}_k \delta_t$. If the time series after the moment of the survey transition is adjusted, then δ_t is defined by (4). If the time series before the changeover is adjusted, then δ_t is defined as

$$\delta_t = \begin{cases} t - T_R & \text{if} \quad t < T_R \\ 0 & \text{if} \quad t \geq T_R \end{cases}. \tag{8}$$

Since the observed series and the estimated discontinuities obey the required consistencies, the adjusted series does too.

4.2 Logratio transformations

The multivariate model developed for $\hat{\mathbf{y}}_t$ accounts for the restriction that $\sum_{k=1}^{K} \hat{y}_{t,k} = 1$, but ignores the restriction $0 \leq \hat{y}_{t,k} \leq 1$. Ignoring the second restriction might result in adjusted parameter estimates taking values outside the admissible range [0,1]. In fact each parameter defines a set of time series that are observed on the *K*-dimensional simplex. One way to account for both restrictions is to apply a logratio transformation to the original data:

$$\hat{x}_{t,k} = \ln\left(\frac{\hat{y}_{t,k}}{\hat{y}_{t,K}}\right), \; k=1, \ldots, K\text{-}1. \tag{9}$$

With (9) the original observations $\hat{\mathbf{y}}_t$ are transformed from the (*K*-1) dimensional simplex to the (*K*-1) dimensional real space, see Aitchison (1986) for details. Instead of modelling the original series $\hat{\mathbf{y}}_t$ and explicitly benchmark the regression coefficients to restriction (7f), it is also possible to develop a set of *K*-1 univariate structural time series models or a set of *K*-1 seemingly unrelated structural time series for $\hat{\mathbf{x}}_t = (\hat{x}_{t,1}, \ldots, \hat{x}_{t,K-1})^t$. This model is obtained with formulae (5) and (6) where $\hat{\mathbf{y}}_t$ is replaced by $\hat{\mathbf{x}}_t$, and taking

$$
\begin{aligned}
\boldsymbol{\alpha}_t &= (L_{t,1}, R_{t,1}, \ldots, L_{t,K-1}, R_{t,K-1}, \beta_1, \ldots, \beta_{K-1})^T, \\
\mathbf{Z}_t &= \left(\mathbf{I}_{[K-1]} \otimes (1,0) \mid \delta_t \mathbf{I}_{[K-1]}\right), \\
\mathbf{T} &= Blockdiag(\mathbf{T}_{tr}, \mathbf{T}_{iv}), \qquad \mathbf{T}_{tr} = \mathbf{I}_{[K-1]} \otimes \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \ \mathbf{T}_{iv} = \mathbf{I}_{[K-1]}, \\
\boldsymbol{\varepsilon}_t &= (\varepsilon_{t,1}, \ldots, \varepsilon_{t,K-1})^T, \\
\boldsymbol{\eta}_t &= (0, \eta_{t,R,1}, \ldots, 0, \eta_{t,R,K-1}, \mathbf{0}_{[K-1]}^T)^T.
\end{aligned}
\tag{10}
$$

The estimated discontinuities apply to the K-1 transformed series. In the case of level intervention, the series observed after the survey transition can be adjusted to the level of the series before the changeover using $\tilde{x}_{t,k} = \hat{x}_{t,k} - \hat{\beta}_k$. The series observed before the survey transition can be adjusted to level under the new situation with $\tilde{x}_{t,k} = \hat{x}_{t,k} + \hat{\beta}_k$. In the case of a slope intervention the time series is adjusted with $\tilde{x}_{t,k} = \hat{x}_{t,k} - \hat{\beta}_k \delta_t$. If the time series after the moment of the survey transition is adjusted, then δ_t defined by (4). If the time series before the changeover is adjusted, then δ_t is defined by (8). Subsequently, the adjusted series can be transformed back to their original values that specify the proportions over K categories on the simplex by the inverse of (9), which is given by

$$
\tilde{y}_{t,k} = \frac{\exp(\tilde{x}_{t,k})}{\sum_{k=1}^{K-1} \exp(\tilde{x}_{t,k}) + 1}, \ k=1, \ldots, K\text{-}1, \tag{11}
$$

$$
\tilde{y}_{t,K} = \frac{1}{\sum_{k=1}^{K-1} \exp(\tilde{x}_{t,k}) + 1}.
$$

The adjusted series meets the consistency property that the adjusted proportions add up to 1, and the values of the K categories take values in the range [0,1], since the logratio transformation accounts for the properties of the data observed on a simplex. Most important drawback of this approach is that the interpretation of the results is more difficult and that the results are not invariant for the choice of the reference category that is used as the denominator in the logratio transformation (9). This implies that the outcomes for the corrected series depend on the choice of the category that is used in the denominator of the logratio transformation.

The asymmetric treatment of the K classes in logratio transformation (9) can be avoided by replacing the reference category $\hat{y}_{t,K}$ in the denominator by the geometric mean over the K categories. This results in the so called central logratio transformation, which is defined by

$$
\hat{z}_{t,k} = \ln\left(\frac{\hat{y}_{t,k}}{g(\hat{y}_t)}\right), \ k=1, \ldots, K, \tag{12}
$$

with

$$g(\hat{y}_t) = \left(\prod_{k=1}^{K} \hat{y}_{t,k} \right)^{\frac{1}{K}} . \tag{13}$$

The advantage of this transformation is that the results do not depend on the choice of a reference category. With (12), however, the vector $\hat{\mathbf{y}}_t$ is transformed from the K-1 dimensional simplex to a linear subspace space of the K dimensional real space that is confined by $\sum_{k=1}^{K} \hat{z}_{t,k} = 0$.

The central logratio transformed series can be modelled with a K dimensional structural time series model. Since the K regression coefficients of the intervention variables must still obey the restriction $\sum_{k=1}^{K} \beta_k = 0$, time series model (5), (6), (7a), through (7f), can be applied to model the series obtained after the central logratio transformation. The series can be adjusted for the estimated discontinuities in a similar way as described for the untransformed and logratio transformed series. Subsequently the adjusted series can be transformed back to their original values by the inverse of (12):

$$\tilde{y}_{t,k} = \frac{\exp(\tilde{z}_{t,k})}{\sum_{k=1}^{K} \exp(\tilde{z}_{t,k})}, \; k=1,\ldots, K. \tag{14}$$

4.3 Benchmarking with series for subpopulations

In sample surveys, parameter estimates for the total population are often also itemized in different subpopulations or domains. The following relationship applies between the series at the national level and its breakdown in H subpopulations

$$\hat{y}_t = \sum_{h=1}^{H} \frac{N_h}{N} \hat{y}_t^H . \tag{15}$$

Here $\hat{y}_t^h$ and N_h denote the parameter estimate and the size of subpopulation h respectively, and $N = \sum_{h=1}^{H} N_h$ the size of the total population. Applying the time series models, described sections 4.1 and 4.2, separately to the series at the national level and its breakdown for these H subpopulations might result in inconsistencies between these series after adjustment for the discontinuities. These inconsistencies arise since the regression coefficients for the intervention variables do not account for the consistency requirement specified by (15).

One solution is to benchmark the adjusted series for the subpopulations to the adjusted series at the national level, for example by using the method of Lagrange multipliers. Let $\tilde{\mathbf{y}}_t = (\tilde{\mathbf{y}}_{t,tot}^T, \tilde{\mathbf{y}}_{t,1}^T \ldots, \tilde{\mathbf{y}}_{t,H}^T)^T$ denote a $(H+1)K$-vector containing the adjusted parameter estimates for period t for the total population $\tilde{\mathbf{y}}_{t,tot} = (\tilde{y}_{t,tot,1}, \ldots, \tilde{y}_{t,tot,K})^T$ and the H subpopulations $\tilde{\mathbf{y}}_{t,h} = (\tilde{y}_{t,h,1}, \ldots, \tilde{y}_{t,h,K})^T$. These parameters must obey a set of linear restrictions such that (15) is met and the unit sum constraint for the vectors $\tilde{\mathbf{y}}_{t,tot}$ and $\tilde{\mathbf{y}}_{t,h}$ for $h = 1, \ldots, H$, still applies. This gives rise to a set of $(H+K)$ linear restrictions that can be expressed as

$$\mathbf{R}\tilde{\mathbf{y}}_t^* = \mathbf{c}, \tag{16}$$

with

$$\mathbf{R} = \begin{pmatrix} (1, -\mathbf{f}_{[H]}^T) \otimes \mathbf{L} \\ \mathbf{I}_{[H+1]} \otimes \mathbf{1}_{[K]}^T \end{pmatrix}, \quad \mathbf{L} = \begin{pmatrix} \mathbf{I}_{[K-1]} & \mathbf{0}_{[K-1]} \end{pmatrix}, \quad \mathbf{f} = \left(\frac{N_1}{N}, \ldots, \frac{N_H}{N} \right)^T, \quad \text{and}$$

$$\mathbf{c} = \left(\mathbf{0}_{[K-1]}^T, \mathbf{1}_{[H+1]}^T \right)^T.$$

Applying the method of Lagrange multipliers gives

$$\tilde{\mathbf{y}}_t^* = \tilde{\mathbf{y}}_t + \mathbf{V}\mathbf{R}^T (\mathbf{R}\mathbf{V}\mathbf{R}^T)^{-1} [\mathbf{c} - \mathbf{R}\tilde{\mathbf{y}}_t], \tag{17}$$

where $\mathbf{V}$ denotes the covariance matrix of $\tilde{\mathbf{y}}_t$. In (17) the discrepancies $[\mathbf{c} - \mathbf{R}\tilde{\mathbf{y}}_t]$ are distributed over the values of $\tilde{\mathbf{y}}_t$ proportional to their accuracy measure specified by $\mathbf{V}$. This implies that the parameters for the total population receive smaller adjustments than the parameters for the subpopulations, since parameters for the total population are estimated more precisely compared to domain estimates. The covariance matrix of (17) is given by

$$\mathbf{V}(\tilde{\mathbf{y}}_t^*) = \mathbf{V} - \mathbf{V}\mathbf{R}^T (\mathbf{R}\mathbf{V}\mathbf{R}^T)^{-1} \mathbf{R}\mathbf{V}.$$

The benchmarked estimates obtained with (17) have smaller variances than the separately adjusted series. The interpretation of this variance reduction is that the restrictions specified by (16) add additional information to the model that is applied to adjust the series for the observed discontinuities.

Inconsistencies can also be avoided by modelling the untransformed series for the total population and its break down in the H subpopulations, i.e. $\tilde{\mathbf{y}}_t = (\tilde{\mathbf{y}}_{t,tot}^T, \tilde{\mathbf{y}}_{t,1}^T, \ldots, \tilde{\mathbf{y}}_{t,H}^T)^T$, simultaneously in one multivariate model and include the consistency requirements in the transition equation for the regression coefficient of the intervention variables. To avoid unnecessary mathematical notation, the transition equation is only given for the regression coefficients of these intervention variables. The formulation of the complete state space representation follows directly from the models defined in section 4.1.

Let $\boldsymbol{\beta} = \mathbf{T}\boldsymbol{\beta}$ denote the transition equation for the time invariant regression coefficients of the intervention variables for the series of the total population and the H subpopulations, i.e. $\boldsymbol{\beta} = \left(\boldsymbol{\beta}_{Tot}^T, \boldsymbol{\beta}_1^T, \ldots, \boldsymbol{\beta}_H^T \right)^T$, with $\boldsymbol{\beta}_{Tot}$ the K dimensional vector containing the intervention variables for the K categories of the parameter for the total population and $\boldsymbol{\beta}_h$ the K dimensional vector containing the intervention variables of the parameter for the subpopulations. If the transition matrix is defined as

$$\mathbf{T} = \begin{pmatrix} \mathbf{O}_{[K\times K]} & \mathbf{f}_{[H]}^{T} \otimes \mathbf{T}_{iv} \\ \mathbf{1}_{[H]} \otimes \mathbf{O}_{[K\times K]} & \mathbf{I}_{[H]} \otimes \mathbf{T}_{iv} \end{pmatrix},$$

where $\mathbf{T}_{iv}$ is defined by (7.f), then it follows that the adjusted series meet the consistencies specified by (15) as well as the unit sum constraint for the *K* classes of the parameter for the total population and the *H* subpopulations.

Both methods can be generalized to benchmark the series for the population total and two or more domain classifications simultaneously. Adding too many restrictions, however, might result in numerical problems for solving (17) or estimating the state space model.

5. Results

The time series models developed in section 4 are applied to the series of separating chemical waste and contact frequency with neighbours, which are plotted in Figure 1 and 2. The results obtained with four different models are compared. These models assume that the series can be decomposed in a stochastic trend, a level intervention and an irregular term. Because the series concern annual data, it was not necessary to use a seasonal component. This allowed the selection of very parsimonious models, which was inevitable since the series are very short (11 years). Adding AR or MA components deteriorated the model fits and generally resulted in overfitting of the data.

The first model is a univariate model, defined by equations (3), (5), (6), and (7a) through (7d), that is applied to each of the series of the separate categories of both parameters. The second model is the multivariate model defined by equations (3), (5), (6), (7a), (7b), (7e), and (7f). This implies that the observed series are not transformed and that the regression coefficients of the intervention variables are explicitly benchmarked by restriction $\mathbf{T}_{iv}$ defined in (7f). The third model is defined by (3), (5), (6) and (10) for the set of *K*-1 series that is obtained after applying the logratio transformation (9). In this case the last category is used as the reference category. The fourth model is defined by equations (3), (5), (6), (7a), (7b), (7e), and (7f) after applying the central logratio transformation (12).

The analysis was conducted with software developed in Ox in combination with the subroutines of SsfPack (beta 3.0) (Doornik 1998 and Koopman *et al*. 1999). Point estimates and standard errors for the regression coefficients of the intervention variable are based on the smoothed Kalman filter estimates using the fixed interval smoother.

For each model two analyses are conducted. One is based on the data available up to and including 2006, the other on the complete series, including 2007. This gives some intuition of the size of the revision of the estimate of the discontinuity if an additional observation under the new approach becomes available.

Estimation results for the discontinuities under the different models are given in Table 1 for the parameter “Separating chemical waste”, and in Table 2 for the parameter “Contact frequency with neighbours”.

As expected in advance, the estimated discontinuities under the univariate model do not obey the restriction $\sum_{k=1}^{K} \hat{\beta}_k = 0$. As a result, the corrected series are not consistent, since the categories for a parameter do not add up to one. This hampers the application of the univariate model.

The multivariate model for the original series and the central logratio transformed series results in consistent series since the estimates for the discontinuities are forced to obey the required restriction. Augmenting the model with restriction (7f) also reduces the standard errors of the estimated discontinuities, since the restriction adds additional information to the model. This follows if the results obtained with the multivariate model for the original series are compared with the results for the univariate model for the original series.

Another way to preserve the consistency between the series of the K categories of a parameter is to apply the logratio transformation, since this transformation eliminates the redundancy due to the unit sum constraint over the K categories. The estimated discontinuities in Tables 1 and 2 refer to the transformed data.

Table 1: Estimated discontinuities for Separating chemical waste with different models

Model	**Category**									
	1		2		3		4		5	
Univ/not transf ‘06	4.29	(1.21)	-4.34	(1.21)	0.00	(1.21)	1.50	(1.21)	-1.44	(1.21)
Univ/not transf ‘07	1.91	(1.88)	-4.15	(0.77)	-0.07	(0.77)	1.49	(0.77)	-1.17	(0.98)
Mvar/not transf ‘06	4.29	(1.07)	-4.35	(1.07)	-0.01	(1.07)	1.50	(1.07)	-1.44	(1.07)
Mvar/not transf ‘07	3.07	(1.44)	-4.01	(0.75)	0.07	(0.75)	1.63	(0.75)	-0.76	(0.98)
Logr transf ‘06 *	-0.06	(0.14)	-1.08	(0.20)	0.16	(0.10)	1.00	(0.20)		
Logr transf ’07 *	0.19	(0.15)	-0.77	(0.21)	0.23	(0.11)	0.68	(0.12)		
C. logr transf ’06*	-0.04	(0.26)	-1.06	(0.26)	0.22	(0.31)	1.01	(0.16)	-0.13	(0.07)
C. logr transf ‘07*	-0.05	(0.25)	-1.09	(0.26)	0.17	(0.30)	1.00	(0.21)	-0.03	(0.07)

**: Results refer to the (central) logratio transformed series. Standard errors in brackets.*

Table 2: Estimated discontinuities for Contact frequency neighbours with different models

Model	**Category**							
	1		2		3		4	
Univ/not transf '06	4.79	(1.19)	0.31	(0.69)	-4.19	(1.32)	1.60	(0.51)
Univ/not transf '07	4.40	(1.20)	-0.09	(0.59)	-3.18	(1.30)	-1.36	(0.59)
Mvar/not transf '06	5.02	(0.93)	0.46	(0.66)	-3.92	(0.96)	-1.56	(0.48)
Mvar/not transf '07	4.44	(0.93)	-0.07	(0.56)	-3.01	(0.95)	-1.35	(0.56)
Logr transf '06 *	0.33	(0.09)	0.27	(0.09)	0.16	(0.09)		
Logr transf '07 *	0.38	(0.11)	0.30	(0.10)	0.14	(0.08)		
C. logr transf '06*	0.14	(0.06)	0.08	(0.06)	-0.03	(0.06)	-0.19	(0.06)
C. logr transf '07*	0.12	(0.05)	0.07	(0.05)	-0.03	(0.05)	-0.16	(0.05)

**: Results refer to the (central) logratio transformed series. Standard errors in brackets.*

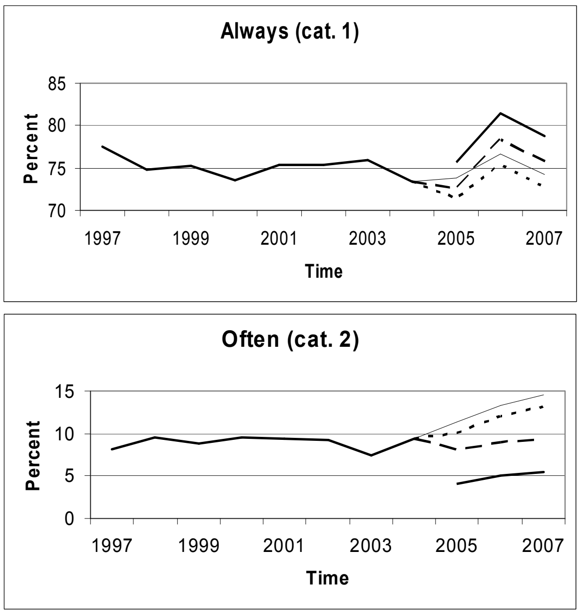
The results obtained under equivalent models illustrate the size of the revision for the estimated discontinuities if the data for an additional year becomes available. Adding the estimates obtained in 2007 to the series results in a revision of the estimated discontinuities. Large revisions are observed for the first category of "Separating chemical waste" under the univariate model and the fourth category of "Contact frequency with neighbours" under the univariate model. For the other three models the sizes of the revisions are smaller with respect to the standard errors. It can be expected that the size of the revisions decreases if the length of the series increases, particularly if the number of data points after the changeover increases.

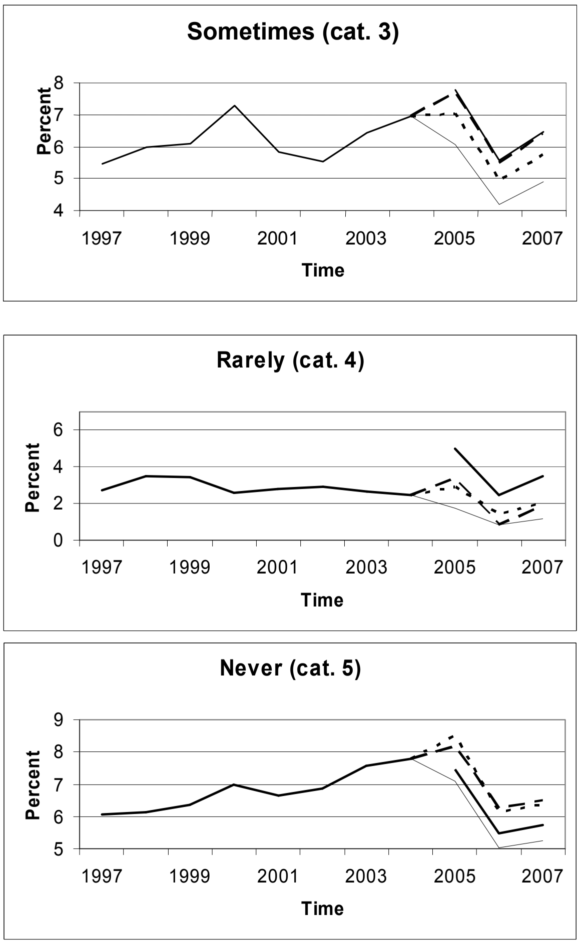
The original data, the corrected series obtained with the multivariate model applied to the original data, the multivariate model applied to central logratio transformed data and the model applied to the logratio transformed data, are shown in figures 3 and 4. The outcomes obtained under the SSPEC for the period 2005 through 2007 are corrected to make the series comparable with the outcomes of the PSLC, using the procedure described in section 4. It is not clear which model results in the most reliable corrections. This requires additional information about the real development of the parameters from sources other than the PSLC and the SSPEC.

The main advantage of the logratio and central logratio transformation is that the adjusted values add up to one and always take values within the admissible range of [0,1] by definition. The major drawback of both transformations is that the interpretation of the results is complex. The estimated discontinuities as well as the corrected series for a particular class are influenced by the discontinuity of the reference class in the case of the logratio transformation. In the case of the central logratio transformation the situation is even more complex, since the estimated discontinuities as well as the corrected series for each particular class is influenced by the discontinuities of all other classes, via the geometric mean over all classes in the denominator of this transformation. Indeed, the results obtained with the central logratio transformation for some categories in Figure 3 (separating chemical waste) look suspect. The estimated discontinuity for the first category is small and not significant (last row Table 1). Nevertheless, there is a substantial deviation between the original series and the series adjusted with the central logratio transformation for the first category. For this category the corrected series appears to be reasonable in line with the series observed before the changeover. The estimated discontinuity for the last category is also very small and not significant (last row Table 1). For this category the series is slightly adjusted in the wrong direction. For the second and the third category it seems that the series are overcorrected in the right direction. The corrected series obtained with the central logratio transformation for the categories of contact with neighbours in Figures 4, on the other hand, are reasonable in line with the series observed before the changeover. An additional disadvantage of the logratio transformation is that the results depend on the choice of the reference category to be used in the denominator of the logratio transformation.

The advantage of the multivariate model applied to the untransformed data is that the interpretation of the results is simple and that the estimated discontinuities for the separated categories are only affected by the other categories through the zero sum constraint. The major drawback is that the corrected values might take values outside the admissible range of [0,1]. This, however, did not occur in this application.

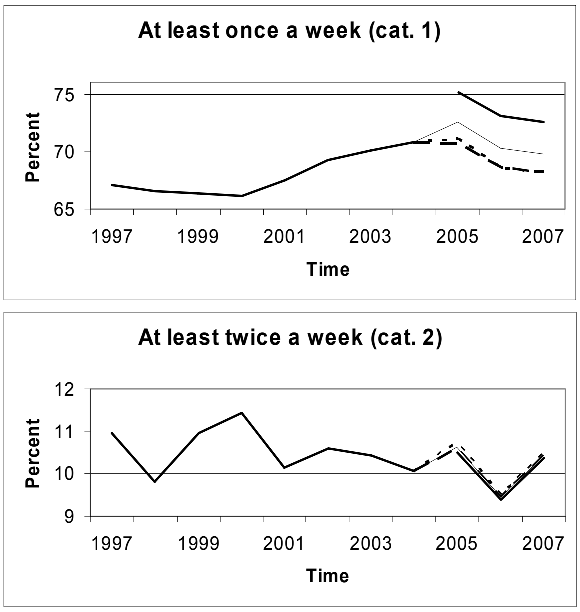
Figure 3: Separating chemical waste

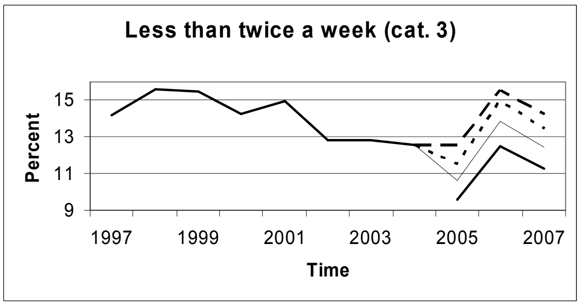


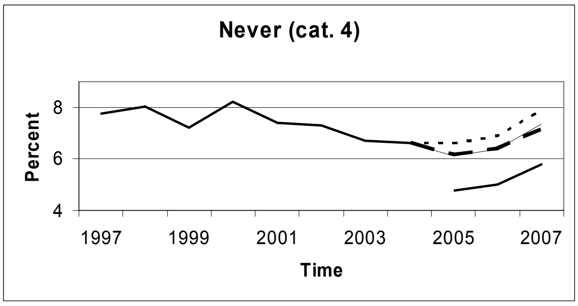


Solid line 1997-2004 estimate based on the PSLC, solid line 2005-2007 estimate based on the SSPEC, dotted line corrected series based on a logratio transformation, dashed line corrected series based on untransformed data, thin solid line corrected series based on central logratio transformation.

Figure 4: Contact frequency with neighbours





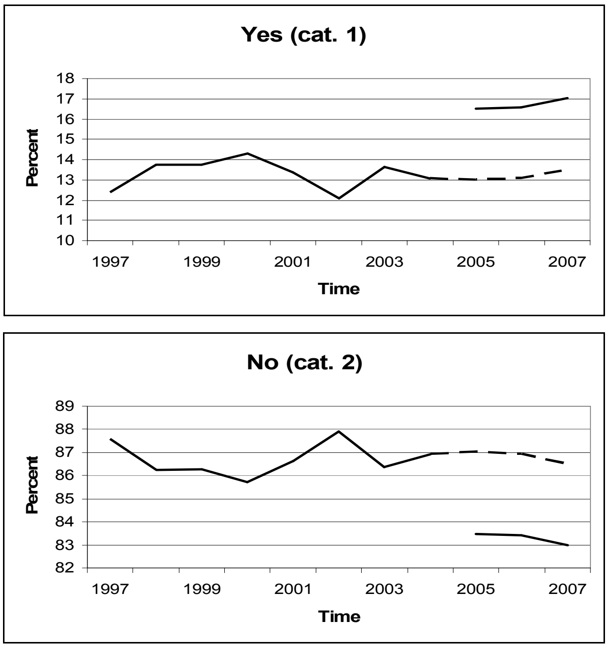


Solid line 1997-2004 estimate based on the PSLC, solid line 2005-2007 estimate based on the SSPEC, dotted line corrected series based on a logratio transformation, dashed line corrected series based on untransformed data, thin solid line corrected series based on central logratio transformation.

Based on these considerations, the multivariate model applied to the untransformed data is finally used in this application to estimate discontinuities and calculate corrected time series for all other parameters about environmental consciousness and social participation. Recall that the differences specified in Annex A are the net effect of real developments, the redesign of the survey, and sample errors. Under the assumed time series model, the net effect of all factors that changed simultaneously in the survey redesign is separated from the real developments and the sampling errors. The estimates for these discontinuities are shown in Annex B. The best estimate for the real development, under the assumed time series model, can be derived from the trend level (not shown).

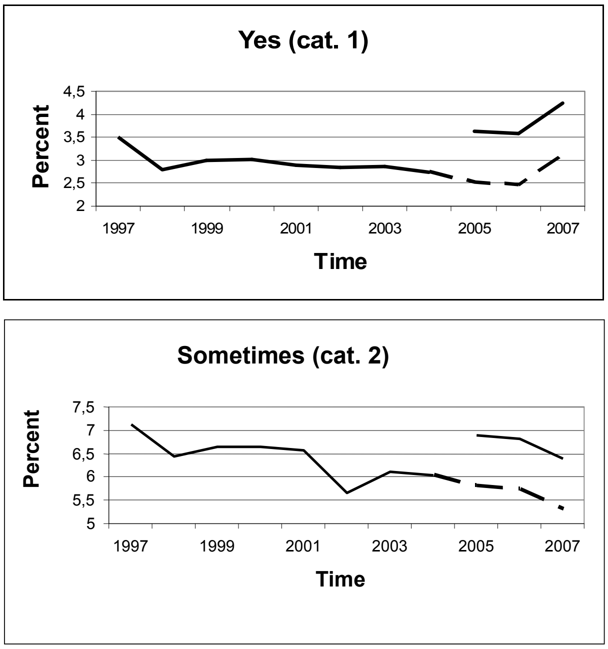
For the parameters sports club activities, suffer from factory smell, and member of circle of friends the estimated series and the corrected series are plotted in Figures 5, 6, and 7 respectively. The outcomes obtained under the SSPEC are corrected by subtracting the estimated discontinuity from the initial estimate to make the series comparable with the outcomes of the PSLC. The leaps in the uncorrected series in the year of the survey transition largely disappear after correction for the estimated discontinuities. Therefore it appears that the corrected series show a more reasonable development of the parameters. There is, however, no hard evidence for this conclusion.

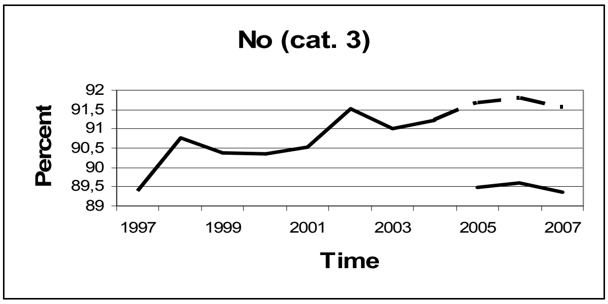
Figure 5: Sports club activities



Solid line 1997-2004 estimate based on the PSLC, solid line 2005-2007 estimate based on the SSPEC, dashed line corrected series based on untransformed data.

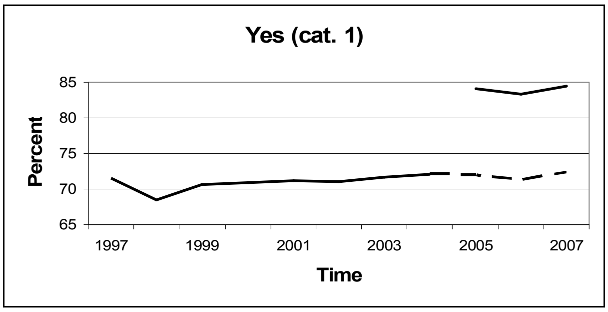
Figure 6: Suffer from factory smell

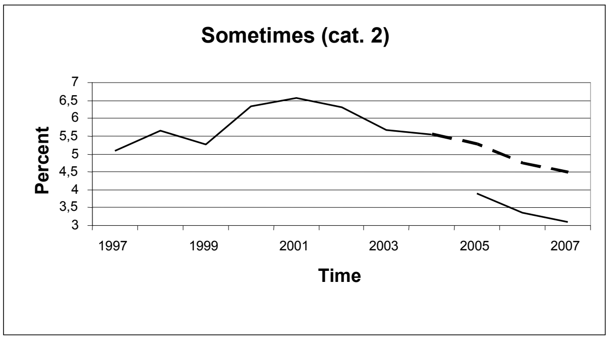


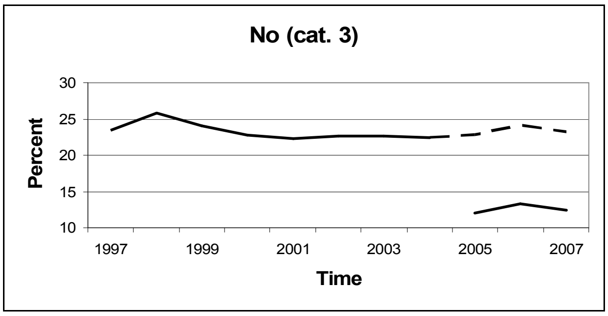


Solid line 1997-2004 estimate based on the PSLC, solid line 2005-2007 estimate based on the SSPEC, dashed line corrected series based on untransformed data.

Figure 7: Member of circle of friends







Solid line 1997-2004 estimate based on the PSLC, solid line 2005-2007 estimate based on the SSPEC, dashed line corrected series based on untransformed data.

In this application, the series for the two domains of gender were also analyzed and adjusted for the observed discontinuities. For a few parameters, the method of Lagrange multipliers, describes in section 4.3, was applied to restore the consistency with the series for the total population. In this case the covariance matrix in (17) was taken diagonal with the variances of the smoothed Kalman filter estimates for the regression coefficients of the intervention variables as elements. This benchmark resulted in small modifications of the adjusted series (results are not shown).

6. Discussions

The transition of the PSLC to the SSPEC resulted in systematic differences in the estimates for parameters about environmental consciousness and social participation. To avoid the confounding of real developments with the systematic effect induced by the redesign on the series of official statistics, structural time series models are developed to estimate the size of the discontinuities. The discontinuity in the series is modelled with an intervention variable that describes the moment that the survey process is redesigned. This approach assumes that the other components of the time series model approximate the real development of the parameter reasonably well and that there is no structural change in e.g. the trend or the seasonal component at the moment that the new survey is implemented. If a change in the real development of the parameter exactly coincides with the implementation of the new survey, then the model will wrongly assign this effect to the intervention variable which is intended to describe the redesign effect. Information available from series of correlated variables can be used to evaluate the assumption that there is no structural change in the real evolution of the parameter. Such auxiliary series can also be added as a regression component to the model with the purpose to reduce the risk that a structural change in the evolution of the series of the target parameter is wrongly assigned to the intervention variable. This risk can also be avoided by conducting an experiment where both surveys are run in parallel for some period of time. This approach is not always feasible due to budget constraints, as was the case for the transition from the PSLC to the SSPEC.

Consistent time series can be obtained by correcting the observed series for the estimated discontinuity. Depending on the anticipated impact of the redesign on the quality of the estimates, the series observed in the past can be adjusted to make it comparable with the outcomes obtained under the new design. It is also possible to adjust the outcomes obtained under the new approach to make them comparable with the series under the old survey design. In this application the data collection mode changed from CAPI under the PSLC to CATI under the SSPEC. Therefore it is anticipated that the series observed in the past are more accurate than the outcomes obtained under the SSPEC. Indeed, with the CAPI mode the entire target population is reached while the CATI mode only surveys the subpopulation with a non-secret telephone number. Furthermore less measurement errors and social desirable answers are expected under the CAPI mode due to the personal contact with an interviewer and the lower interview speed, see e.g. Holbrook et al. 2003, Roberts 2007. Based on these considerations it was decided that the outcomes obtained under the SSPEC are corrected to make the series comparable with the outcomes of the PSLC. Under the assumption that the development observed with the CATI data is representative for the entire target population, consistent time series are obtained.

One aspect of the time series approach is that more observations under the new approach become available when time proceeds. The advantage is that the discontinuities can be quantified more accurately if this additional information becomes available. A concomitant drawback is that the estimated discontinuities three years after redesigning the survey are still subject to revisions. A publication policy is required to deal with these revisions in practice. For the moment it is decided to base the final estimates for the discontinuities on the information available up until 2007 in this application. If the data for 2008 result in a substantial revision, then it will be considered to revise the estimates again.

The common picture of the effect of the redesign is an increase of the proportion of respondents in the first categories compensated by a decrease in the last categories (see Annex B). This is the net effect of all factors that changed simultaneously in the changeover from the PSLC to the SSPEC (see section 3).

In general the categories of the parameters about social participation are ordered from strong civic and socially participated behaviour in the first category to civic and social exclusion in the last category (see section 2.3). Therefore it appears that under the CATI mode, a higher level of civic and social participation is observed. Although the effects of all factors that changed are confounded, there are some indications that this is mainly caused by a selection effect. First, the order of the answer categories of the parameter “feeling isolated” and “social contacts” is reversed, that is the first category corresponds with social exclusion while the last category corresponds with social inclusion. For these parameters, the effect of the redesign is also reversed, that is a decrease of the proportion of respondents in the first category compensated by an increase in the last category. This is a strong indication that the observed effect is a selection effect and not a mode effect. Furthermore, this conclusion is in line with the finding of Van den Brakel et al. (2008), where a strong decrease is observed of the fraction of respondents that is satisfied with police performance during their last contact with the police, when the data collection changed from CATI to a mixed mode design with CAPI and CATI.

For the parameters about environmental consciousness, and particularly the parameters about suffer of smell and noise, it is less clear that the observed increase of the proportion in the first category at the cost of a decrease in the latter categories is caused by a selection effect. Probably mode effects or context effects play a more prominent role for these parameters. The SSPEC is introduced as a survey that is focussed on topics about social participation and environmental consciousness while the PSLC is introduced as a more general survey on living conditions. It might be expected that a higher level of smell and noise burden is reported under the SSPEC, since the survey is directly focussed on these kinds of topics.

References

Aitchison, J. (1986). *The statistical analysis of compositional data*. Chapman and Hall: London.

Brackstone, G. (1999). Managing data quality in a statistical agency. *Survey Methodology*, **25**, 139-149.

Cochran, W.G. (1977). *Sampling Techniques*, third edition, John Wiley & Sons, New York.

De Leeuw, E. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, **21**, 233–255.

Dillman, D.A. & Christian, L.M. (2005). Survey mode as a source of instability in responses across surveys. *Field Methods*, **17**, 30-52.

Doornik, J.A. (1998). *Object-oriented matrix programming using Ox 2.0*. London: Timberlake Consultants Press.

Durbin, J. & Koopman, S.J. (2001). *Time series analysis by state space methods*. Oxford: Oxford University Press.

Harvey, A.C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge: Cambridge University Press.

Holbrook, A.L., M.C. Green, and J.A. Krosnick (2003). Telephone versus Face-to-Face Interviewing of National Probability Samples with Long Questionnaires. *Public Opinion Quarterly*, 67, pp. 79-125.

Kalton, G. and Schuman, H. (1982). The effect of the question on survey responses: a review. *Journal of the Royal Statistical Society*, **145** (Series A), 42-73.

Koopman, S.J., Shephard, N. & Doornik, J.A. (1999). Statistical algorithms for models in state space using SsfPack 2.2. *Econometrics Journal*, **2**, 113-166.

Roberts, C. (2007). Mixing modes of data collection in surveys: A methodological review. Review paper, NCRM/008, National Centre for Research Methods, City University London.

Särndal, C-E., B. Swensson and J. Wretman (1992). *Model Assisted Survey Sampling.* New York: Springer Verlag.

Van den Brakel, J.A. (2008). Design-based analysis of embedded experiments with applications in the Dutch Labour Force Survey. *Journal of the Royal Statistical Society*, **171**, (Series A), 581-613.

Van den Brakel, J.A., P.A. Smith and S. Compton (2008). Qualtiy procedures for survey transiotions, experiments, time series and discontinuities. *Journal for Survey Research Methods*, **2**, 123-141.

Annex A

Estimated discontinuities for parameters about environmental consciousness and social participation

Parameter	Categories									
	1		2		3		4		5	
Youth work	3.05**	(0.47)	-3.05**	(0.47)						
School activities	3.03**	(0.61)	-3.03**	(0.61)						
Welfare or nursary work	2.41**	(0.59)	-2.41**	(0.59)						
Member of library	4.10**	(0.97)	-4.10**	(0.97)						
Member of sports club	6.40**	(0.96)	-6.40**	(0.96)						
Member of labour union	2.09**	(0.81)	-2.09**	(0.81)						
Member of hobby club	2.09**	(0.60)	-2.09**	(0.60)						
Sports club activities	3.46**	(0.73)	-3.46**	(0.73)						
Hobby club activities	0.74	(0.44)	-0.74	(0.44)						
Culture club activities	1.49**	(0.48)	-1.49**	(0.48)						
Labour Union activities	0.60	(0.32)	-0.60	(0.32)						
Suffer from factory noise	0.79**	(0.27)	0.49	(0.29)	-1.28**	(0.40)				
Suffer from traffic noise	4.44**	(0.76)	0.63	(0.76)	-5.07**	(0.96)				
Suffer from train noise	1.50**	(0.37)	0.52	(0.40)	-2.02**	(0.53)				
Suffer from (un)loading noise	1.49**	(0.36)	-0.53	(0.38)	-0.95	(0.51)				
Suffer from airplane noise	1.60**	(0.50)	1.47*	(0.68)	-3.08**	(0.80)				
Suffer from factory smell	0.90**	(0.36)	0.85	(0.50)	-1.75**	(0.60)				
Suffer from fireplace smell	1.32**	(0.41)	2.63**	(0.58)	-3.96**	(0.69)				
Suffer from farm smell	1.93**	(0.34)	2.55**	(0.63)	-4.46**	(0.70)				
Suffer from traffic smell	2.80**	(0.40)	0.91*	(0.42)	-3.71**	(0.57)				
Feeling isolated	-0.58	(0.30)	-1.97**	(0.53)	2.56**	(0.60)				
Speaking with people	1.60**	(0.50)	-1.17**	(0.43)	-0.43	(0.27)				
Social contacts	-3.13**	(0.75)	-4.92**	(0.74)	8.06**	(0.95)				
Sympathy of other people	2.87**	(0.64)	-2.30**	(0.57)	-0.57	(0.32)				
Member of circle of friends	12.08**	(0.83)	-1.65**	(0.43)	-10.43**	(0.75)				

Freq. visiting cinema	2.99**	(0.67)	0.72	(0.80)	-0.27	(0.88)	-3.44**	(1.01)		
Freq. visiting concerts and opera	3.22**	(0.44)	7.63**	(0.77)	6.36**	(0.99)	-17.21**	(0.98)		
Freq. visiting dance evenings	8.29**	(0.76)	1.42**	(0.55)	2.15**	(0.62)	-11.85**	(0.96)		
Freq. visiting theatre	0.56*	(0.26)	2.85**	(0.57)	8.98**	(0.95)	-12.39**	(1.00)		
Freq. of trips in countryside	14.79**	(1.00)	-4.56**	(0.78)	-2.03**	(0.66)	-8.20**	(0.70)		
Freq. visiting playgrounds	2.20**	(0.41)	0.36	(0.49)	0.80	(0.74)	-3.36**	(0.89)		
Freq. visiting restaurant	11.75**	(0.98)	-3.24**	(0.95)	-4.70**	(0.80)	-3.81**	(0.62)		
Freq. of mind games	0.84	(0.54)	1.18	(0.80)	-0.76	(0.79)	-1.26	(1.01)		
Freq. of do-it-yourself jobs	4.32**	(0.58)	6.82**	(0.93)	-2.81**	(0.88)	-8.33**	(0.98)		
Freq. of contact with neighb.	4.38**	(0.90)	0.46	(0.62)	-2.99**	(0.63)	-1.84**	(0.47)		
Separating chemical waste	2.26**	(0.89)	-5.25**	(0.50)	0.79	(0.53)	2.54**	(0.39)	-0.33	(0.54)
Separating glass	-0.89	(0.83)	0.60	(0.53)	0.87	(0.49)	0.45	(0.33)	-1.01*	(0.45)
Separating paper waste	1.33	(0.69)	0.65	(0.39)	-0.24	(0.32)	-0.24	(0.17)	-1.51**	(0.50)
Separating vegetable waste	3.20**	(0.98)	-2.18**	(0.51)	-0.49	(0.48)	-0.02	(0.33)	-0.52	(0.84)
Separating garden waste	-2.18**	(0.74)	-1.92**	(0.40)	0.10	(0.31)	0.14	(0.20)	3.88**	(0.56)
Satisfaction with circle of friends	9.37**	(0.76)	0.88	(0.99)	-8.70**	(0.98)	-0.75	(0.47)	-0.81**	(0.27)
Statisfaction with leisure activities	6.44**	(0.62)	4.00**	(0.94)	-7.16**	(1.01)	-1.42*	(0.62)	-1.86**	(0.39)
Bike instead of car for short dist.	2.66**	(0.97)	5.18**	(0.92)	-1.58*	(0.79)	0.06	(0.55)	-6.31**	(0.60)

Standard errors in brackets. Category labels are explained in section 2.

**: p-value <0.05; **:p-value<0.01.*

Annex B

Estimated discontinuities for parameters about environmental consciousness and social participation

Parameter	Categories									
	1		2		3		4		5	
Youth work	3.40**	(0.25)	-3.40**	(0.25)						
School activities	2.55**	(0.33)	-2.55**	(0.33)						
Welfare or nursary work	1.88**	(0.63)	-1.88**	(0.63)						
Member of library	4.13**	(0.57)	-4.13**	(0.57)						
Member of sports club	7.74**	(0.78)	-7.74**	(0.78)						
Member of labour union	2.76**	(0.38)	-2.76**	(0.38)						
Member of hobby club	1.99**	(0.31)	-1.99**	(0.31)						
Sports club activities	3.53**	(0.50)	-3.53**	(0.50)						
Hobby club activities	0.73*	(0.35)	-0.73*	(0.35)						
Culture club activities	1.49**	(0.20)	-1.49**	(0.20)						
Labour Union activities	0.52	(0.32)	-0.52	(0.32)						
Suffer from factory noise	0.69**	(0.21)	0.23	(0.23)	-0.93**	(0.24)				
Suffer from traffic noise	5.25**	(0.65)	0.00	(0.38)	-5.25**	(0.66)				
Suffer from train noise	1.31**	(0.37)	0.65	(0.43)	-1.96**	(0.44)				
Suffer from (un)loading noise	1.57**	(0.26)	-0.73**	(0.19)	-0.85**	(0.26)				
Suffer from airplane noise	2.07**	(0.38)	1.45**	(0.38)	-3.52**	(0.38)				
Suffer from factory smell	1.12**	(0.27)	1.08**	(0.27)	-2.20**	(0.27)				
Suffer from fireplace smell	1.39**	(0.47)	2.10**	(0.59)	-3.49**	(0.61)				
Suffer from farm smell	1.84**	(0.38)	2.71**	(0.66)	-4.56**	(0.68)				
Suffer from traffic smell	2.41**	(0.29)	0.91**	(0.29)	-3.32**	(0.29)				
Feeling isolated	-0.42**	(0.18)	-1.31**	(0.53)	1.73**	(0.53)				
Speaking with people	1.75**	(0.41)	-1.52**	(0.41)	-0.23	(0.41)				
Social contacts	-1.67**	(0.66)	-6.73**	(0.72)	8.41**	(0.66)				
Sympathy of other people	2.58**	(0.55)	-2.05**	(0.54)	-0.53	(0.37)				
Member of circle of friends	12.14**	(0.97)	-1.38**	(0.53)	-10.76**	(0.96)				

Freq. visiting cinema	2.48**	(0.88)	1.45	(0.88)	-0.07	(0.88)	-3.86**	(0.88)		
Freq. visiting concerts and opera	2.96**	(1.21)	8.05**	(1.33)	5.93**	(1.11)	-16.94**	(1.43)		
Freq. visiting dance evenings	8.09**	(0.64)	1.40*	(0.64)	1.49*	(0.64)	-10.98**	(0.64)		
Freq. visiting theatre	0.76**	(0.27)	2.62**	(0.65)	8.59**	(1.09)	-11.97**	(1.11)		
Freq. of trips in countryside	14.84**	(0.77)	-2.57**	(0.91)	-3.17**	(0.86)	-9.10**	(0.77)		
Freq. visiting playgrounds	2.15**	(0.17)	0.14	(0.44)	1.82*	(0.82)	-4.10**	(0.83)		
Freq. visiting restaurant	11.76**	(1.53)	-1.62	(0.96)	-5.76**	(1.45)	-4.38**	(1.31)		
Freq. of mind games	1.38**	(0.51)	0.61	(0.81)	-0.80	(0.73)	-1.19	(0.87)		
Freq. of do-it-yourself jobs	4.55**	(0.58)	6.42**	(0.81)	-2.92**	(0.96)	-8.06**	(1.10)		
Freq. of contact with neighb.	4.44**	(0.93)	-0.07	(0.56)	-3.01**	(0.95)	-1.35**	(0.56)		
Separating chemical waste	2.98*	(1.46)	-4.02**	(0.75)	0.13	(0.78)	1.68*	(0.77)	-0.77	(0.98)
Separating glass	0.13	(0.67)	0.54	(0.40)	0.81*	(0.40)	0.36	(0.40)	-1.84**	(0.40)
Separating paper waste	0.07	(0.75)	1.08*	(0.49)	0.13	(0.43)	-0.37	(0.28)	-0.91	(0.66)
Separating vegetable waste	4.48**	(0.84)	-2.00**	(0.47)	-0.63	(0.48)	-0.34	(0.47)	-1.51*	(0.73)
Separating garden waste	-2.42**	(0.91)	-1.55**	(0.52)	-0.31	(0.36)	-0.04	(0.36)	4.33**	(0.71)
Satisfaction with circle of friends	10.44**	(0.70)	-0.04	(0.49)	-9.38**	(0.76)	-0.10	(0.45)	-0.92**	(0.18)
Statisfaction with leisure activities	7.40**	(0.87)	3.41**	(1.13)	-7.27**	(1.37)	-1.47*	(0.72)	-2.07**	(0.58)
Bike instead of car for short dist.	3.31**	(0.92)	5.27**	(0.93)	-2.17**	(0.69)	0.16	(0.69)	-6.57**	(0.69)

Standard errors in brackets. Category labels are explained in section 2.

**: p-value <0.05; **:p-value<0.01.*