Model-Based Estimation for Official Statistics

08

Jan van den Brakel and Jelke Bethlehem

Discussionpaper (08002)

Statistics Netherlands

Voorburg/Heerlen, 2008

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
Prinses Beatrixlaan 428
2273 XZ Voorburg

second half of 2008:
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
http://www.cbs.nl

ISSN: 1572-0314

6008308002 X-10

Summary:

This paper summarizes the advantages and disadvantages of design-based and model-assisted estimation procedures that are widely applied by most of the European national statistical institutes. Several situations are identified were model-based approaches can have additional value in the production of official statistics, e.g. to deal with small sample sizes, measurement errors and discontinuities due to survey redesigns. It is concluded that there is a case for having official releases and time series, with appropriate quality and methodology descriptions, which rely on statistical models for situations where design-based estimators do not result in sufficiently reliable estimates.

Keywords: Design-based inference, Model-based inference, Non-sampling errors, Small Area Estimation, Survey redesign.

1. Introduction

The purpose of survey sampling is to obtain statistical information about a finite population by selecting a probability sample from this population, measuring the required information about the units in this sample and estimating finite population parameters such as means, totals and ratio's. The statistical inference in this setting can be design-based, model-assisted or model-based. In the design-based and model-assisted approach, the statistical inference is based on the stochastic structure induced by the sampling design. Parameter and variance estimators are derived under the concept of repeatedly drawing samples from a finite population according to the same sampling design, while statistical modelling plays a minor role. This is the traditional approach of survey sampling theory, followed by authors like Hansen, et al. (1953), Kish (1965), Cochran (1977), and Särndal et al., (1992). In the model-based context, the probability structure of the sampling design plays a less pronounced role, since the inference is based on the probability structure of an assumed statistical model. This is the position taken by authors like Gosh and Meeden (1997), Valliant et al. (2000), and Rao (2003). An overview of the different modes of inference in survey sampling is given by Little (2004).

Currently, the application of model-based estimation procedures at Statistics Netherlands is limited. This appeared to be the situation in most European national statistical institutes (NSIs). Several factors are responsible for the slow adoption of these methods. One is the fact that many European NSIs are rather reserved in the application of model-based estimation procedures and generally rely on the more traditional design-based or model-assisted procedures for producing their official statistics. NSIs need to play safe in the production of official statistics and therefore do not want to relay on model assumptions, particularly if they are not verifiable. Another factor is that the methodology of advanced model-based techniques, used e.g. in the context of small area estimation, is intellectually and practically inaccessible (Heady and Ralphs, 2005). Indeed the statistical theory is rather complex and the available software at NSIs is often not suitable to conduct the required calculations in a straightforward manner in a production environment. This hampers the implementation in survey processes to produce timely official statistics.

In this paper situations are explored were model-based procedures can be used to produce official statistics that are more reliable than the more standard design-based approaches. The focus will be mainly on the use of linear mixed models and time series models in the estimation of official releases and not on the use of statistical modelling at a micro level, for example to handle missing data and automatic data editing. The selected topics do not constitute an exhaustive list of situations were statistical modelling has potential applications. They are chosen, since they are currently on the research agenda of Statistics Netherlands's Methodology Department.

In section 2 design-based and model-assisted estimation techniques are reviewed and it is motivated why this approach is widely used by NSIs. In the remaining sections, situation are described were these techniques are less appropriate to produce sufficiently reliable statistics and model-based techniques might be preferable. These are small area estimation (section 3), estimation in the presence of non-sampling errors (section 4), dealing with discontinuities due to survey redesigns (section 5), and the use of register data (section 6). In section 7 other examples were statistical modelling plays an important role in the production of official releases are shortly summarized. The paper concludes with some general remarks in section 8.

2. Design-based and model-assisted estimation

Design-based and model-assisted estimators refer to a class of estimators that expand or weight the observations in the sample with the so-called survey weights. Survey weights are derived from the sampling design and available auxiliary information about the target population. Functions of the expanded observations in the sample are used as (approximately) design-unbiased estimators for the unknown population parameters of interest. The associate inferences are based on the probability distribution induced by the sampling design with the population values held fixed.

A well known design-based estimator is the π-estimator or Horvitz-Thompson estimator, developed by Narain (1951), and Horvitz and Thompson (1952) for unequal probability sampling from finite populations without replacement. The observations are weighted with the inverse of the inclusion probability, also called design-weights. This estimator is design-unbiased, since the expectation of the estimator with respect to the probability distribution induced by the sampling design is equal to the true but unknown population value.

The precision of the Horvitz-Thompson estimator can be improved by making advantage of available auxiliary information about the target population. In the model-assisted approach developed by Särndal et al. (1992) this estimator is derived from a linear regression model that specifies the relationship between the values of a certain target parameter and a set of auxiliary variables for which the totals in the finite target population are known. Based on the assumed relationship between the target variable and the auxiliary variables, a generalized regression estimator can be derived of which most well-known estimators are special cases. After this estimator is derived, it is judged by its design-based properties, such as design expectation and design variance.

Generalized regression estimators are non-linear. Only a linearized approximation, obtained by means of a Taylor series expansion that is truncated at the first order term, can be shown to be design-unbiased. Therefore generalized regression estimators are approximately design-unbiased. Also the design variances of generalized regression estimators are approximated with the variance of this linearized approximation. Generalized regression estimators are, nevertheless, asymptotically design-unbiased and consistent, see Isaki and Fuller (1982), and Robinson and Särndal (1983). In the context of finite population sampling, asymptotic results are obtained in a framework of an infinite sequence of increasing populations such that the sample size and the population size both tend to infinity.

Generalized regression estimators are members of a larger class of calibration estimators, Deville and Särndal (1992). Callibration estimators minimally adjust the design-weights under a pre-specified loss function such that the sum over the weighted auxiliary variables in the sample adds up to the known population totals. Under a quadratic loss function, the generalized regression estimator is obtained as a special case. Early papers of Luery (1986) and Alexander (1987) anticipated on the more complete treatment of calibration estimation by Deville and Särndal (1992).

Due to the following properties, generalized regression estimators are widely applied by NSIs in the production of official statistics. The (approximate) design unbiasedness and consistency of generalized regression estimators is an attractive and important property since it provides a form of robustness in the case of large sample sizes. If the underlying linear model of the generalized regression estimator explains the variation of the target parameter in the finite population reasonably well, then this might result in a reduction of the design variance of the Horvitz-Thompson estimator. If the model is misspecified, then this might result in an increase of the design variance but the property that the generalized regression estimator is approximately design unbiased remains. From this point of view, the generalized regression estimator is robust against model-misspecification.

Although auxiliary information was originally used in the design and estimation procedure of a survey to decrease the sampling variance of estimators, nowadays it is an important tool to decrease the bias due to selective non-response. Estimators using auxiliary information are generally more robust against selective non-response than estimators that do not use auxiliary information, Särndal and Swenson (1987), Bethlehem (1988), and Särndal and Lundström (2005).

Another important property is that the generalized regression estimator ensures that for the auxiliary variables the weighted observations sum up to the known population totals. This property is used to enforce consistency between the marginal totals of different publication tables.

Besides accuracy, timeliness is also an important quality aspect for official statistics. In the daily practice of survey sampling, the generalized regression estimator is often used to produce one set of weights for the estimation of all target parameters of a multi-purpose sample survey. As pointed out above, the generalized regression estimator is robust in the sense that model-misspecification does not compromise design-consistency. For these reasons, this estimator is very attractive to produce timely official releases in a regular production environment.

3. Small Area Estimation

Design-unbiasedness is a useful property for large sample sizes, but it is often incompatible with reliable estimates for small sample sizes. A major drawback of generalized regression estimators are the relatively large design variances in the case of small sample sizes. If sample sizes are too small to apply direct survey estimators and additional information is available, model-dependent estimation procedures might be used to produce sufficiently reliable statistics. In these procedures a model is applied to borrow information from other related data sets to improve the precision of the estimates. Compared to the traditional design-based survey estimators, model-based estimators have much smaller variances. The price that is paid for this variance reduction is that these model-based estimators are more or less design-biased. The size of the bias depends on the correctness of the model. Model-misspecification easily results in severely biased estimates. Careful model selection and validation is therefore a central part of the application of these model-dependent procedures. The estimation of domain parameters for which insufficient data are available to apply design-based or model-assisted procedures is the realm of small area estimation (SAE), see Rao (2003) for an overview.

Small sample sizes arise if estimates are required for very detailed geographic or socio-demographic classifications. Small sample sizes also arise if timely estimates are required, resulting in a lack of time to collect a sufficient amount of data. Another important application for small area estimation is the production of timely short-term economic indicators. Since there is a strong demand for these timely indicators, many NSIs work with provisional releases that are based on the data obtained in the first part of the data collection period. From this point of view SAE can improve the timeliness of a survey process.

Many SAE procedures use sample information observed in other domains. In this context, domains refer to geographic or socio demographic classifications of the target population. This is generally referred to as borrowing strength over space. The common approach is to allow for random domain effects in a linear mixed model and apply a composite estimator using methods like empirical best linear unbiased prediction (EBLUP), empirical Bayes, or hierarchical Bayes. These estimators can be considered as a weighted average of a design-based estimator and a synthetic regression type estimator, see Rao (2003) for a comprehensive overview. These approaches are currently investigated in a research project to estimate annual municipal unemployment figures in the Dutch Labour Force Survey (LFS), see Boonstra et al. (2007).

Most surveys conducted by NSIs operate continuously in time and are based on cross-sectional or rotating panel designs. Consequently, SAE procedures that borrow strength from data collected in the past as well as cross sectional data from other small areas are particularly interesting. The unemployment rate in month t, for example, will be correlated with the unemployment rate in the preceding periods. Since the LFS is conducted continuously in time, it is efficient to improve the estimates for month t with the sample information observed in preceding periods.

The common approach to borrow strength in time and space is to allow for random area and random time effects in a linear mixed model and apply a composite estimator like the BLUP or EBLUP. Rao and Yu (1994) extended the linear mixed model with an AR(1) model to combine cross-sectional data with information observed in preceding periods. See also EURAREA (2004) for linear mixed models that allow for spatial and temporal autocorrelation in the random terms, with the purpose to improve the precision of small domain estimators with sample information obtained in other periods and domains.

A different approach is followed by Pfeffermann (1991), Pfeffermann and Burck (1990), Pfeffermann and Bleuer (1993), and Pfeffermann and Tiller (2006). They combine time series data with cross-sectional data by modeling the correlation between the parameters of the time series models of the separate domains in a multivariate structural time series model.

The estimation procedure based on structural time series models for repeated surveys has high practical value. This approach borrows strength over time and space and can be made more robust against model-misspecification by benchmarking the sum of the small area estimates to the direct estimates at an aggregated level. This property provides a “built-in mechanism” against model-misspecification. Furthermore it is possible to specify models that explicitly account for the rotating panel design of the survey and the autocorrelation between the different panels, resulting in more efficient estimates for the population parameters. It can also be extended to account for rotation group bias (see section 4). First results to apply this approach to estimate monthly unemployment rates for six demographic domains in the Dutch LFS are described by Krieg and Van den Brakel (2007).

This time series approach also fits into a framework for producing timely short-term statistics. At the start of a data-collection period, the model yields forecasts for the population parameters for time periods for which no survey data are available (this is sometimes called nowcasting). When new survey data become available, timely preliminary and final estimates can be produced, taking advantage from data collected in the past and in neighboring areas. This results in a smooth conversion from predicted values, provisional releases to final releases.

4. Estimation in the presence of non-sampling errors

The traditional design-based approach of survey sampling assumes that the observations obtained from the sampling units are true fixed values observed without error. Stochasticity is introduced since a probability sample is drawn from a finite target population. Expectations and variances of the estimators are derived with respect to the probability distribution induced by the sampling design with the population values held fixed. It is emphasized in section 2 that auxiliary information is used in the generalized regression estimator to correct, at least partially, for the bias due to selective non response. Nevertheless design-based estimation procedures generally do not handle non-sampling errors in an effective way.

One example where measurement errors result in problems with data integrity is the rotating panel design of the Dutch LFS. Each month a sample of addresses is drawn and data are collected by means of computer assisted personal interviewing of the households selected in the sample. The sampled households are re-interviewed by telephone four times at quarterly intervals. Based on the generalized regression estimator, each month estimates about the employment and unemployment are obtained for the preceding three months.

A major problem with this panel design is that systematic differences occur between the subsequent waves due to mode effects, panel effects and panel attrition. This is a well known problem for rotating panel designs, and is in the literature referred to as rotation group bias (RGB), see Bailar (1975). An analysis conducted by Van den Brakel and Krieg (2007a, 2007b) illustrates that the RGB in the monthly unemployment rate induced by the rotating panel design result in substantial differences in the trends and the seasonal patterns between the subsequent waves. This RGB clearly illustrates the existence of non-sampling errors such as measurement errors and panel attrition. Therefore the traditional concepts that observations obtained from sampling units are true fixed values observed without error and that the respondents can be considered as a representative probability sample from the target population, generally assumed in design-based sampling theory, is not tenable under such designs. The application of direct estimators in the case of measurement errors and selective panel attrition will result in severely biased estimates.

An additional problem is that the monthly sample size of the Dutch LFS is too small to rely on the generalized regression estimator to produce official statistics about the monthly employment and unemployment (see section 3). To deal with the problems of small sample sizes and the RGB an estimation procedure based on multivariate structural time series modeling is applied to the monthly data of the Dutch LFS, Van den Brakel and Krieg (2007a, 2007b). With this time series model a substantial increase of the accuracy of the monthly estimates for the unemployment rate is obtained. Firstly, the model explicitly estimates the RGB in the trend and the seasonal effects between the first wave and the four subsequent waves. As a result, estimates for the unemployment rates are corrected for this RGB. Secondly, the time series model borrows strength from data observed in preceding periods via the assumed model for the population parameter and the autocorrelation between the survey errors of the different panels.

The time series model yields estimates for the trend and seasonal components of the population parameter. Seasonally adjusted parameter estimates and their estimation errors are therefore obtained as a by-product of this estimation procedure. Another major advantage is that this approach accounts for the autocorrelation in the survey errors due to the rotating panel design. Pfeffermann et al. (1998) show that ignoring these autocorrelation, for example with the Henderson filters in X11-ARIMA, results in spurious trend estimates.

5. Dealing with discontinuities due to survey transitions

Many surveys run by NSIs are continuous, and a significant aspect of their value comes from that continuity, sometimes over very long periods. Methods, procedures and definitions applied in the survey gradually become outdated, which makes change and improvement inevitable from time to time. This, however, may affect the continuity of the time series. Therefore it is important to minimise the impact of such changes, to keep inconvenience for users to a minimum.

It is well known that adjustments in the survey process can affect response bias and therefore the parameter estimates of a sample survey. When an ongoing survey is changed, it is not clear whether a change in the series is a result of a real development or is induced by the redesign. Even if no change in the series is observed, it is still possible that a real development could be nullified by an opposite redesign effect.

In an ideal transition process, it is avoided that the autonomous development is confounded with redesign effects. In cases where the data remain the same, the differences can be investigated by recalculation, for example the introduction of new editing, imputation or estimation methods. Also a new economic activity classification system in business surveys results in discontinuities in time series, which can be quantified using the same data with the addition of the new classification. Where data collection or coding procedures are affected, however, the data are not consistent, and a natural way to evaluate the effect of the change in the new approach is to conduct a field experiment where the regular and new survey designs are run concurrently for some period. This allows us to estimate the main survey parameters under both survey designs and test whether these estimates are significantly different. Due to e.g. resource constraints it will not always be possible to conduct experiments to quantify the effect of a redesign. In such situations an intervention analysis through time series modelling might be a second best alternative to estimate the effect of the redesign on the outcomes of a survey.

There are several ways to deal with observed discontinuities. A conservative approach is to quantify the discontinuities only for the period in which both approaches are run concurrently (without extrapolation). This implies that the autonomous development in the series is separated from the effect of the redesign on the parameter estimates for this period only. This can be considered as a design-based and rather safe approach since the observed effects are not extrapolated beyond the period where both approaches were run concurrently. On the other hand, this generally does not meet the users' requirements, since they often desire uninterrupted series for e.g. policy evaluation.

Other methods, which meet the requirement of maintaining uninterrupted series, rely on a model to adjust the series for the observed difference beyond the period where both approaches are run in parallel. Different time series models and synthetic approaches are available to adjust the series observed in the past under the regular design, to make them comparable with the figures obtained under the new design, and these are discussed in Van den Brakel, Smith and Compton (2007). This is sometimes referred to as backcasting. These approaches are applied to the series about Police performance and Crime Victimization to deal with discontinuities due to the redesign of the Permanent Survey on Living Conditions and the Population Police monitor which are in 2005 integrated into one new survey; the Security Monitor.

For NSIs it might be preferable to quantify the effect of a redesign and separate this effect from autonomous developments through experiments and avoid any extrapolation of observed differences. There is, on the other hand, a case for having one official series, adjusted for this type of discontinuities rather than allowing each user to generate a different version of the series for their own use. Since the NSIs collect the data and have access to the micro data, they should be capable to produce the most reliable consistent series. This fits in the policy of Statistics Netherlands, since the division of Macro-economic Statistics and Dissemination has an expert centre for long time series, with the purpose to produce long uninterrupted series.

6. Use of register data

There is a persistent pressure that NSIs must reduce costs and response burden for businesses and organizations by replacing survey data for register data. These trends make model-based procedures more and more attractive and relevant for NSIs to apply in the production of official statistics, Chambers et al. (2006).

For the short-term economic indicators, for example, the possibilities to collect survey data in the strata of the small and medium sized enterprises are more and more restricted, since it is expected that Statistics Netherlands compiles these statistics from the Value Added Tax (VAT) registrations.

Replacing survey data for register data in this particular application will, however, give rise to several problems. A substantial part of the population, for example, declares their VAT quarterly or even annually. Since most short term economic indicators are produced on a monthly basis, model-based estimation techniques are a realistic alternative to produce monthly official releases for the subpopulations that do not declare VAT on a monthly basis. The VAT register data for the subpopulations that declare on a monthly and a quarterly basis is more or less complete after about 45 days after the month or quarter respectively. As a result, the timeliness of the statistics, compiled with these register data is another point of concern. As pointed out in section 3, estimation procedures that are based on time series modelling can be used to produce preliminary timely estimates. Finally model-based procedures can be used as temporary back-up scenarios in cases were registrations are delayed or suddenly fall away.

Note, however, that the use of register data instead of survey data is not quite a new development and can also be compatible with the design-based and model-assisted approach discussed in section 2. Statistics Netherlands' Regional Income Survey, for example, is since 1946 based on a probability sample and the data are completely obtained from tax registrations and registrations about social benefits. Regional income distributions are produced with design-based estimators like the generalized regression estimator, see Van den Brakel en De Jong (1994), and Van den Brakel and Nieuwenbroek (1996). Wallgren and Wallgren (2007) provide a comprehensive overview how official releases can be compiled from register data.

7. Other applications

The four situations were statistical modelling can play an important role for official statistics, discussed in the preceding sections, is far from exhaustive. There are, of course, many more cases were statistical modelling plays a crucial role in the statistical process. An important example is the macro integration techniques for the national accounts. In practice, the different sources used to compile the national accounts are made consistent using subject matter knowledge. Such approaches heavily rely on implicit, naive model assumptions. There are, however, more formal approaches available, which are based on explicit modelling of the available macro data. See e.g. Denton (1971), Sefton and Weale (1995), and Magnus et al. (2000) for a variety of model-based procedures to adjust national accounts. The potentials of these techniques are currently investigated in the redesign of the national accounts.

Statistical modelling plays an important role in the production of hedonic price indices, Balk (2008). These approaches are applied at Statistics Netherlands to compile the price index of new dwellings.

Time series modelling is frequently applied to produce seasonally adjusted official releases. In the soft-packages TRAMO-SEATS, ARIMA models play a prominent role to model and eliminate the seasonal component in a series, see Gomez and Maravall (2000). In the software-package X12 ARIMA, which is used at Statistics

Netherlands, the seasonal components in a series are subtracted from the series through a battery of linear filters, see Findley et al. (1998). To apply the symmetric filters of this package up to the most recent observation in the series, it is required that the observed series is extended with forecasts. Since these forecasts are obtained through ARIMA modelling, the entire approach followed with X12 ARIMA is principally model-based.

Another example is the estimation in a mixed-mode data collection setting. Here, measurement errors can result in problems with the data integrity and model-based approaches might be more effective to handle measurement errors compared to design-based approaches, see Cobben et al. (2007).

So far, the discussion is focussed on the use of statistical methods in the estimation of population parameters, i.e. at a macro level. Statistical modelling also plays a role on a micro level to handle missing data through imputation. See e.g. Little and Rubin (1987) for a comprehensive overview. See Tempelman (2007) and Pannekoek and De Waal (2005) for applications to Statistics Netherlands' business surveys.

8. Discussion

Statistics Netherlands is generally rather reserved in the application of model-based estimation procedures for the production of official statistics. The most important reasons are that design-based approaches are robust against model-misspecification and are convenient to apply in a production environment to produce timely official statistics. In this paper several areas are mentioned were the model-based approach can have an additional value to design-based and model-assisted approach. There is a growing demand for detailed statistics. Examples are monthly unemployment and annual municipal unemployment figures. In these situations direct estimators have unacceptable large standard errors and model-based estimators are required to produce more accurate statistics.

Measurement errors can result in severe problems with the integrity of the collected data, which cannot be handled in a satisfactory way with standard design-based approaches. One example is the bias between the subsequent waves in the rotating panel design of the Dutch LFS. A model-based approach, based on a multivariate structural time series model, can be used to produce estimates that account for this bias and make efficient use of the sample information obtained in preceding periods.

Survey redesigns and survey transitions generally affect the continuity of series. Model-based approaches can be useful to quantify the effect of a redesign and to meet the users requirement of having uninterrupted series.

Model-based procedures can be used in situations were register data are used to compile official statistics and problems arise with timeliness, coverage, and continuity of the available register data.

It can be concluded that there are cases for having official series that are based on model-based procedures for situations where direct estimators do not result in sufficiently reliable estimates. Such releases should be clearly separated from the design-based releases, and accompanied with appropriate methodology and quality descriptions, were the underlying model assumptions are made explicit.

References

Alexander, C.H. (1987). A Class of Methods for Using Person Controls in Household Weighting. *Survey Methodology*, 13, pp. 183-198.

Balk, B.M. (2008). *Price and Quantity Index Numbers; Models for Measuring Aggregate Change and Difference*, in press. New York: Cambridge University Press.

Bailar, B.A. (1975). The Effects of Rotation Group Bias on Estimates from Panel Surveys. *Journal of the American Statistical Association*, 70, pp. 23-30.

Bethlehem, J.G. (1988). Reduction of nonresponse bias through regression estimation. *Journal of Official Statistics*, 4, pp. 251-260.

Boonstra, H.J., B. Buelens, and M. Smeets (2007). Estimation of Municipal Unemployment Fractions – A Simulation Study Comparing Different Small Area Estimators. Unpublished research paper, BPA nr. DMK-DMH-2007-04-20-HBTA, Statistics Netherlands, Heerlen.

Brakel, J. A. van den, en W. de Jong 1994. Het schatten van de steekproefomvang voor de RIS 1994. Interne CBS nota, BPA no. 12775-94-RSM, Sector Statistische Methoden, Heerlen.

Brakel, J.A. van den, en N. Nieuwenbroek 1996. Ophoging van het Regionale Inkomensonderzoek. Interne CBS nota, BPA no.: 5070-96-RSM, Sector Statistische Methoden, Heerlen.

Brakel, J.A. van den, and S. Krieg (2007a). Estimation of the Monthly Unemployment Rate through Structural Time Series Modelling in Rotating Panel Design. Discussion paper; forthcoming , Statistics Netherlands.

Brakel, J.A. van den, and S. Krieg (2007b). Modelling Rotation Group Bias and Survey Errors in the Dutch Labour Force Survey. *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

Brakel, J.A. van den, P. Smith, and S. Compton (2007). Quality procedures for survey transitions – experiments and discontinuities. Discussion paper 07005, Statistics Netherlands and the Office for National Statistics, UK.

Chambers, R., J.A. van den Brakel, D. Hedlin, R. Lehtonen, and Li-Chun Zhang (2006). Future Challenges of Small Area Estimation. *Statistics in Transition*, 7, pp. 759-769.

Cobben, F., B. Janssen, K. van Berkel, and J.A. van den Brakel (2007). Statistical Inference in a Mixed-Mode Data Collection Setting. Invited paper for the ISI conference, Lissabon, 2007.

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

Denton, F.T. (1971). Adjustment of monthly or quarterly series to annual totals: an approach based on quadratic minimization. *Journal of the American Statistical Association*, 66, pp. 99-102.

Deville, J., and C.-E. Särnal (1992). Calibration Estimators in Survey Sampling, *Journal of the American Statistical Association*, 87, pp. 376-382.

EURAREA (2004). Project reference volume, deliverable D7.1.4, Technical report, EURAREA consortium.

Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C. & Chen, B.C. (1998). New capabilities and methods of the X-12-ARIMA Seasonal Adjustment Program. *Journal of Business and Economic Statistics*, 16, pp. 127-176 (with Discussion).

Gomez, V. & Maravall, A. (2000). Automatic modeling methods for univariate series. In D. Pena, G.C. Tao & R.S. Tsay (Eds.) *A Course in Time Series* (chapter 7). New York: Wiley .

Gosh, M., and G. Meeden (1997). *Bayesian Methods for Finite Population Sampling*. London: Chapman & Hall.

Heady, P., and M. Ralphs, (2005). EURAREA: an overview of the project and its findings. *Statistics in Transition* 7, 557—570.

Hansen, M.H., W.N. Hurwitz, and W.G. Madow (1953). *Sample Survey Methods and Theory*, Vol I and II. New York : Wiley.

Horvitz, D.G., and D.J. Thompson (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Ass*ociation, 47, pp. 663-685.

Isaki, C.T., and W.A. Fuller (1982). Survey design under the regression superpopulation model. *Journal of the American Statistical Association*, 77, pp. 89-96.

Kish, L. (1965). *Survey Sampling*. New York: Wiley.

Krieg, S., and J.A. van den Brakel (2007). Model evaluation for multivariate structural time series models for the Dutch Labour Force Survey. *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

Little, R.J.A. (2004). To Model or Not to Model? Competing Modes of Inference for Finite Population Sampling, *Journal of the Royal Statistical Association*, 99, pp. 546-556.

Little, R.J.A., and D.B. Rubin (1987). *Statistical Analysis with Missing Data*. New York: Wiley.

Luery, D.M. (1986). Weighting Sample Data Under Linear Constraints on the Weights, *Proceedings of the section on Social Statistics*, American Statistical Association, pp. 325-330.

Magnus, J.R, J.W. van Tongeren, and A.F. de Vos (2000). National Accounts Estimation Using Indicator Ratios, *The Review of Income and Wealth* 46, pp. 329-350.

Narain, R. (1951). On sampling without replacement with varying probabilities. *Journal of the Indian Society of Agricultural Statistics*, 3, pp. 169-174.

Pannekoek, J., and T. De Waal (2005). Automatic Edit and Imputation for Business Surveys; The Dutch Contribution to the EUREDIT Project. *Journal of Official Statistics*, 21, pp. 257-286.

Pfeffermann, D. (1991). Estimation and Seasonal Adjustment of Population Means Using Data from Repeated Surveys. *Journal of Business & Economic Statistics*, 9, pp. 163-175.

Pfeffermann, D., and S.R. Bleuer (1993). Robust Joint Modelling of Labour Force Series of Small Areas. *Survey Methodology*, 19, pp. 149-163.

Pfeffermann, D., and L. Burck (1990). Robust Small Area Estimation combining Time Series and Cross-sectional Data. *Survey Methodology*, 16, pp. 217-237.

Pfeffermann, D., M. Feder, and D. Signorelli (1998). Estimation of Autocorrelations of Survey Errors with Application to Trend Estimation in Small Areas. *Journal of Business & Economic Statistics*, 16, pp. 339-348.

Pfeffermann, D., and R. Tiller (2006). Small Area Estimation with State Space Models Subject to Benchmark Constraints. *Journal of the American Statistical Association*, 101, pp. 1387-1897.

Rao, J.N.K. (2003). *Small Area Estimation*. New York: Wiley.

Rao, J.N.K., and M. Yu (1994). Small-area estimation by combining time-sereis and cross-sectional data. *The Canadian Journal of Statistics*, 22, pp. 511-528.

Robinson, P.M., and C.-E. Särndal (1983). Asymptotic properties of the generalized regression estimator in probability sampling. *Sankhya B*, 45, pp 240-248.

Särndal, C.-E., and S. Lundström (2005). Estimation in Surveys with Nonresponse. New-York: Wiley.

Särndal, C.E., and B. Swensson (1987). A general view of estimation for two phases of selection with application to two-phase sampling and nonresponse. *International Statistical Review*, 55, pp. 279-294.

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

Sefton, J., and M.R. Weale (1995). *Reconciliation of national income and expenditure.* Cambridge: Cambridge University Press.

Tempelman, C. (2007). *Imputation of restricted data; applications to business surveys*. Ph.D. Thesis; Groningen: Groniningen University.

Valliant, R., A.H. Dorfman, and R.M. Royall (2000). *Finite Population Sampling and Inference, A Prediction Approach.* New York: Wiley.

Wallgren and Wallgren (2007). *Register-based Statistics: Administrative Data for Statistical Purposes*. New York: Wiley.