Models in official statistics

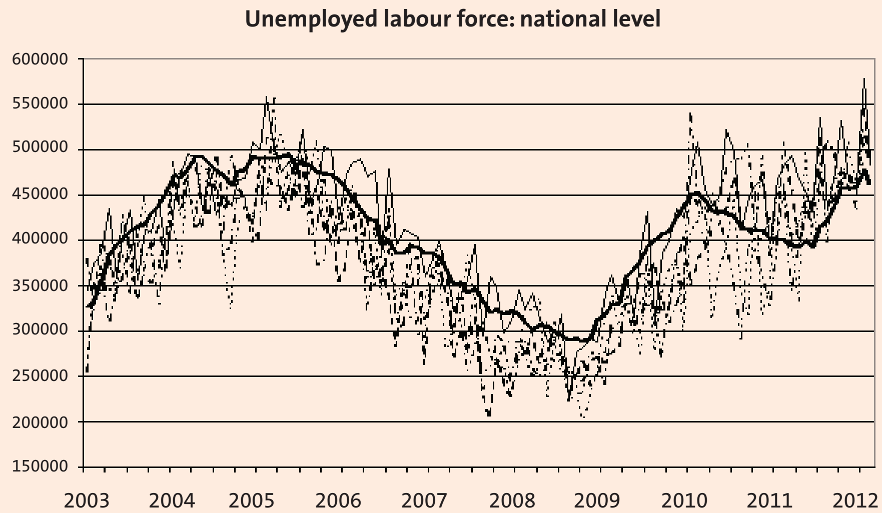
Maastricht University

Inaugural lecture

Jan A. van den Brakel

School of Business and Economics

Models in official statistics



Models in official statistics

Colofon

Ontwerp en print: Océ Business Services, Maastricht

ISBN: 978-90-5681-388-8
NUR: 870

Models in official statistics

Inaugural lecture
Delivered by accepting the position of extraordinary professor of Survey Methodology at the Maastricht University School of Business and Economics

Maastricht, 27 April 2012

Jan A. van den Brakel

Rector magnificus, dean, board of the School of Business and Economics, colleagues from the Maastricht University and Statistics Netherlands, family and friends. Welcome and thank you for coming to Maastricht to listen to this inaugural speech.

Familie, vrienden en bekenden, van harte welkom en mijn dank om naar Maastricht te komen om mijn inaugurele rede bij te wonen.

Liebe Freunde und Bekannte aus Gangelt, herzlich willkommen in Maastricht und vielen Dank, dass Ihr an meiner Antrittsvorlesung teilnehmt.

Last week, I a read an article titled: 24% of people older than 40 suffer from asthmatic disorders. 24%, that is a lot. This statement was based on a sample of people from Maastricht. Apparently this sample is representative of a larger population of people in general, at least with respect to asthma prevalence. But is this true? If you know that Maastricht is located in a valley that frequently suffers from smog and high concentrations of air pollution, than it is clear that the generalisation made by the author is hard to defend. Survey methodology – which is the title of my chair – refers to the statistical science that should prevent researches for making this kind of erroneous statements.

An inauguration is an introduction of a new professor to the public and to the University. In this speech I will explain the purpose of my chair, which is installed by Statistics Netherlands and the Maastricht University School in Business and Economics and positioned within the Department of Quantitative Economics. My aim with this talk is to provoke your interest in the field of survey methodology and official statistics.

Introduction

National statistical institutes, like Statistics Netherlands, are mandated by law to publish statistical information about economic and social developments of a society. This information is often referred to as official statistics. The required data are obtained via registrations or collected through surveys, usually on the basis of a sample.

Survey methodology studies the statistical theory and methodology that is required to produce this kind of reliable statistical information about modern societies using information that is available from registrations and survey samples. Statistical inference is the methodology used to make statements about these unknown variables. The statistical inference applied in this context can be design based, model assisted or model based. Design based and model assisted means that the inference is based on the probability structure of the sample design that is used to draw a sample from the target population. Model based means that the inference is based on a statistical model that describes how a random variable is related to one or more other random variables according to an assumed probability distribution.

In this talk, I will explain the differences between these three modes of inference. I will start with the design-based and model-assisted approach. They are traditionally used by national statistical institutes. They are popular since they do not rely on an explicitly assumed statistical model. For decades, there has been the prevailing opinion that official statistics must be free from model assumptions, since model misspecification easily translates into wrong statements about the variable of interest. I will continue with a discussion about the price that is paid by following this opinion. Three situations are presented where the use of model-based methods has made valuable contributions to the more traditional design-based and model-assisted approaches. I will finish this talk with a plea for using model-based procedures in official statistics.

Design-based inference

Consider a finite population that contains N elements, for example the 17 million people currently residing in the Netherlands. Our interest is focussed on relevant statistical information about this population. This information is often defined as totals, means or proportions. An example of an important figure, published by Statistics Netherlands, is the total number of unemployed people in the Netherlands. This information is not only required at national level but also for all kinds of subpopulations, like municipalities, age classes or gender classes.

The population values for these variables are generally unknown. Until the beginning of the twentieth century this kind of information was obtained by a complete census of the target population. This implies that the variable of each element in the population is measured. It is clear that this is very laborious and expensive. Therefore the concept of random sampling has been developed, mainly on the basis of the work of Neyman (1934) as a method of obtaining valid estimators for finite population parameters based on a modest but representative sample, rather than on a complete census. Other important milestone papers are Hansen and Hurwitz (1943), Narain (1951), and Horvitz and Thompson (1952).

The selection of the sample is based on a probability mechanism, which ensures that the sample is representative of the target population. An estimator of the unknown population total is obtained as the sum over the observations in the sample, expanded with the so called design weights. These weights are constructed such that the sum over the weighted observations is an unbiased estimate of the unknown population total. Under the design-based approach, these weights are derived from the sampling design and are obtained as the inverse of the probability that a sampling unit is included in the sample. For example, if a person has the probability of one in a hundred to be included in the sample, then the design weight is equal to a hundred. This implies that this observation represents hundred units in the population. In sampling theory this is a well known estimator and is called the Horvitz-Thompson estimator.

This estimation procedure is called design based, since inference is completely based on the randomization distribution induced by the sampling design. Statistical modelling of the observations obtained in the survey does not play any role so far. At this point survey sampling is almost unique in statistical science, with the exception of Kempthorne (1955) where a randomization approach for the analysis of randomized experiments is proposed in a way that is similar to the design-based inference approach in sampling theory.

Design-based inference is a very powerful concept that is still used in modern statistical science because:

1) It allows drawing valid inference of unknown variables of a large population based on a relatively small but representative sample.
2) Uncertainty of using an estimator of the unknown population total can be measured by calculating the design variance of this estimator.
3) The precision of the estimator can be improved by taking advantage of auxiliary information in the design of the sample. Examples are stratified sampling with optimal allocation and sampling designs where selection probabilities are approximately proportional to the target variable.

Model-assisted inference

The second inference mode I want to discuss is model-assisted estimation. National statistical institutes often have auxiliary information about the target population from external sources. An example is the distribution of people over age classes and regions which is known from municipal registrations. The precision of the Horvitz-Thompson estimator can be improved by making advantage of this auxiliary information. One way is to improve the efficiency of the sampling design, as discussed before. Another way is to use this auxiliary information in the estimation procedure via the so called general regression estimator proposed by Särndal et al. (1992).

As in the case of the Horvitz-Thompson estimator, the general regression estimator expands the observation in the sample with a regression weight such that the sum over the weighted observations is an approximately design-unbiased estimator of the unknown population total. The design weights of the Horvitz-Thompson estimator are adjusted such that the sum over the weighted auxiliary variables in the sample equates to the known population totals. This results in a correction for groups that are underrepresented in the sample, for example due to selective nonresponse.

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 variable and a set of auxiliary variables for which the totals in the finite target population are known. Most estimators known from sampling theory can be derived as a special case from the general regression estimator. Examples are the ratio estimator and poststratification.

General regression estimators are members of a larger class of calibration estimators, Deville and Särndal (1992). Calibration 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 general 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).

The general regression estimator has two very attractive properties. Although this estimator is derived from a linear model, it is still approximately design unbiased. If the underlying linear model explains the variation of the target parameter in the population reasonably well, then the use of this auxiliary information will result in a reduction of the design variance compared to the Horvitz-Thompson estimator and it might also decrease the bias due to selective nonresponse, Särndal and Swenson (1987), Bethlehem (1988), and Särndal and Lundström (2005). Model misspecification might result in an increase of the design variance but the property that this estimator is approximately design unbiased remains. From this point of view, the general regression estimator is robust against model misspecification. The linear model is only used to derive an estimator that uses auxiliary information but the resulting estimator is still judged by its design-based properties, such as design expectation and design variance. This is the reason that this approach is called model assisted.

Another attractive property of the general regression estimator is that only one set of weights is required for the estimation of all target parameters of a multi-purpose sample survey. This is not only convenient from a practical point of view, but also enforces consistency between the marginal totals of different publication tables of the survey.

For these two reasons, this estimator is very attractive for producing timely official releases in a regular production environment.

Model-based inference

Results published by national statistical institutes must enjoy public confidence. For decades, this has resulted in the prevailing opinion that methods used to produce official statistics must be free from model assumptions and should therefore be based on the above mentioned design-based and model-assisted approaches. These approaches, however, have some severe limitations. A major drawback is that they have large design variances in the case of small sample sizes and do not handle measurement errors effectively. In such situations model-based estimation procedures can be used to produce more reliable estimates. Model based refers to procedures that rely on the probability structure of an explicitly assumed statistical model, whereas the probability structure of the sampling design plays a less pronounced role.

Examples of situations where model-based procedures can provide valuable contributions in the production of official statistics are:

- Small area estimation.
- Dealing with discontinuities in series of statistics that are induced by a redesign of the survey process.
- The use of alternative data sources.

These topics are discussed in more detail in the remainder of this talk. Other examples were model-based estimation procedures can be applied are inference in mixed-mode data collection procedures to handle measurement errors, Buelens and Van den Brakel (2011), and estimation procedures for informative designs to handle sample designs were the selection or the response mechanism depends on the target variable, Pfeffermann (2011).

Small area estimation

Design-based and model-assisted estimators only use the sample information that is observed in a particular domain and over a specific period. A major drawback of these estimators is that they have unacceptable large design variances in the case of small sample sizes. The term “small area” should not be taken literally: Small areas refer to domains or subpopulations for which the sample size is so small that design-based and model-assisted estimation procedures would result in estimates that are too imprecise. They occur if:

- Estimates are required for detailed breakdowns of the population.
- Estimates are required for relatively short periods.

Variables for a particular domain are often correlated with the same variables from other domains. If the unemployment rate for men in this month increases, then it is very likely that the unemployment rate for women will also increase.

Variables are also correlated with values observed in preceding periods. The unemployment rate in this month will be strongly related to the unemployment rate in preceding months. Many surveys are conducted repeatedly over time. Therefore it is efficient to use sample information observed in other domains or in preceding periods to improve the precision of the domain estimates.

Small area estimation refers to a class of estimation procedures that explicitly rely on statistical models to take advantage of sample information that is observed in other domains or preceding periods, see Rao (2003). Two approaches are identified: multilevel modelling and time series modelling.

1. Multilevel modelling
The first approach is based on multilevel models, Fay and Herriot (1979), Battese et al. (1988). These models are predominantly used to take advantage of cross-sectional sample information that is observed in other domains. They consist of a regression component, where available auxiliary information is used to explain the variation in the survey data, and a random component, which describes the unexplained variation between the domains. Through the regression component, sample information from other domains is used to improve the precision of the estimates for each domain separately.

Consider as an example the situation where unemployment figures at the level of municipalities are estimated by the Dutch Labour Force Survey. About 65,000 observations are obtained in the sample on an annual basis. Sample sizes within the municipalities vary between 10 and 2500. An auxiliary variable is the number of people formally registered as being unemployed, which, at municipal level, is known exactly. Despite the availability of this register, unemployment has to be measured by a survey, since people that are unemployed are not always formally registered.

A design-based estimator only uses the sample information observed in each particular municipality. For many municipalities few observations are available, which is insufficient to make reliable design-based estimates. The regression component of the multilevel model uses the entire sample to estimate the relationship between the target variable, measured with the Labour Force Survey and the auxiliary variable. Since the value for this auxiliary variable is known for each municipality, the regression model can be used to make a precise prediction of the unemployed labour force for each municipality. Through the available auxiliary information, sample information from other domains is used to increase the effective sample size for each separate domain.

Finally, a model-based composite estimator is obtained for the separate domains using methods like empirical best linear unbiased prediction, empirical Bayes and hierarchical Bayes. This can be interpreted as a weighted average of the regression prediction and the design-based estimator. Particularly for the municipalities with a limited number of observations, the precision of this model-based estimator will be much larger than the design-based estimator since it is based on a much larger sample. In this application this approach reduced standard errors in small municipalities by 30%, Boonstra et al. (2008, 2010). The price paid by using this method is that model misspecification results in a design bias that is not reflected in the mean squared errors of this estimator.

2. Time series modelling

The second approach is based on time series modelling. The Labour Force Survey for example is a continuous survey and used to produce reliable monthly estimates about the unemployed labour force. Design-based estimators only use sample information that is observed in a particular month. This is not very efficient, since the unemployed labour force in a particular month is strongly correlated with the values of preceding periods.

As an alternative, for each domain, a structural time series model can be assumed for the series observed with the direct estimator. Under such models sample information observed in the preceding periods can be combined with the current sample estimate to obtain a more precise model-based estimate for the current period. If time series for the different domains were combined in a multivariate time series model, then sample information observed across domains could also be used to improve the precision of the separate domain estimates by modelling the correlation between the disturbances of the time series components across domains. This approach is proposed by Pfeffermann and Burck (1990) and Pfeffermann and Bleuer (1993). In the case of strong correlation between the disturbances of the time series components, more parsimonious common factor models can be formulated to further improve the efficiency of this estimation approach, Krieg and Van den Brakel (2012).

The model can also be extended by series of related auxiliary information. Modelling the correlation between the disturbances of the time series components between the auxiliary series and the series of the target variables, will further improve the precision of the domain estimates for the target variables, Van den Brakel and Krieg (2011).

Example

The monthly sample size of the Dutch Labour Force Survey is considered to be too small to produce reliable monthly figures by the general regression estimator. Therefore a structural time series model is since 2010 used to produce more precise monthly unemployment figures. This method is initially proposed by Pfeffermann (1991).

In Figure 1, the general regression estimates for the monthly unemployed labour force for the domain men with an age between 25 and 44, are compared with the estimates obtained under this time series model. The solid line shows the series of the general regression estimates. The dashed line is the series of the filtered estimates obtained under the time series model. This model is applied to each domain separately. The series of the general regression estimates is more volatile due to the relatively large sampling errors. The series of the model estimates show a smoother pattern since the survey errors are filtered from the series of the general regression estimates.

Figure 1: Comparison general regression estimates (GREG) and filtered estimates (STM regular) under a structural time series model for unemployed labour force for men 25-44.

Figure 2 compares the standard errors of the general regression estimates with the standard errors obtained with three different versions of the time series model. The solid line shows the standard errors of the general regression estimates. The line with the squares shows the standard errors of the filtered estimates obtained with the time series model that is applied to each domain separately and which is currently used to produce official monthly statistics about the labour force. The difference measures the increase in precision by taking advantage of the sample information observed in preceding periods within a domain.

The broken line shows the standard errors of a time series model that models the correlation between the trends of six different domains simultaneously. The difference with the standard errors of the regular time series model – that is the line with the squares – measures the increase in precision by modelling the sample information observed in other domains.

The line with the triangles shows the standard errors of the regular time series model applied to each domain separately, extended by an auxiliary series of the number of people that are formally registered as unemployed. The difference with the standard errors of the regular time series model – that is the line with the squares – measures the increase in precision by modelling the correlation with strongly related auxiliary information.

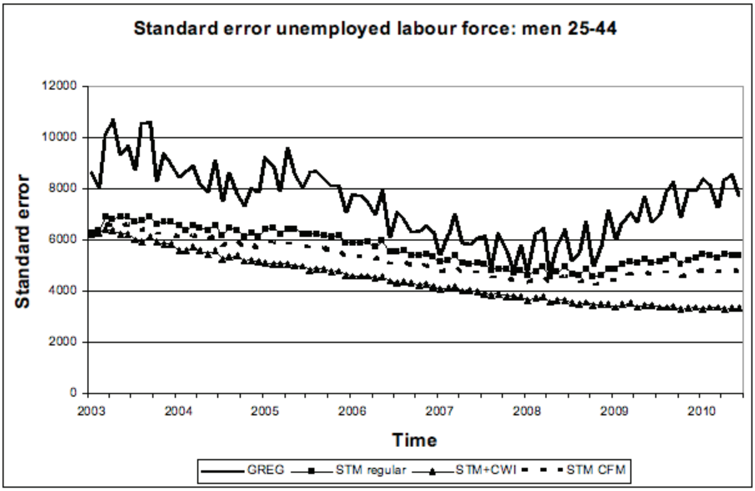


Figure 2: Comparison standard errors general regression estimates (GREG) and filtered estimates unemployed labour force for men 25-44 under three different time series model; the regular model applied to each domain separately (STM regular), the regular model extended with an auxiliary series of people formally register as unemployed (STM+CWI), and a common factor model for all domains (STM CFM).

Discontinuities
The second example of model-based inference deals with discontinuities due to survey redesigns. Sample surveys conducted by national statistical institutes are generally conducted repeatedly over time. This results in a series of statistics that describe the evolution of population variables of interest. A significant aspect of their value comes from the comparability of the outcomes over time. Modifications and redesigns of the underlying survey process generally have a systematic effect on the outcomes of a sample survey. Therefore survey processes are generally kept unchanged as long as possible with the purpose to maintain uninterrupted series. Methods and procedures become gradually outdated. Therefore we have to redesign the survey process itself from time to time. To avoid confounding real developments with systematic effects that are induced by the redesign, it is important to quantify these discontinuities.

There are several ways to quantify such discontinuities. A straightforward approach is to conduct the old and new design in parallel for some period of time through a large scale field experiment. This allows the analyses of systematic differences between design-based estimates obtained under both approaches; see e.g. Van den Brakel (2008). Generally, we need a large sample to accurately observe pre-specified differences. Often this is not possible due to budget constraints.

Therefore, in many situations a parallel run is not available or only at an insufficient sample size. In such cases alternative methods, which are based on explicit statistical models, should be considered to quantify the effect of a redesign.

If a parallel run is missing, then the evolution of the variable of interest can be modeled with an appropriate structural time series model. An intervention variable that describes the moment of the change-over from the old to the new survey process is added to disentangle the systematic effect induced by the redesign of the survey from real developments of the variables of interest, Van den Brakel et al. (2008) and Van den Brakel and Roels (2010).

In many cases there is a limited budget for a parallel run. In such cases, the regular survey used for official publication purposes is conducted in full scale while the alternative approach is conducted with a limited sample size. This is in fact the intermediate case of the two aforementioned situations. The sample size allocated to the regular survey will be sufficiently large to apply design-based estimators, at least for the planned domains. The sample size allocated to the alternative approach is, on the other hand, not sufficiently large and the aforementioned small area estimators can be considered to obtain sufficiently precise model-based estimates. This is an interesting application. Besides the auxiliary information which is available from censuses and registrations, there are also adequate direct estimates for the same variables available from the regular survey, Van den Brakel et al. (2012). These variables are, however, subject to sampling errors and small area estimators that use this information as auxiliary variables must account for this uncertainty, Ybarra and Lohr (2008).

Example
This example further elaborates on the application of the time series model for the estimation of monthly labour force data, described under small area estimation. Before we focus on the effects of a redesign of the LFS, I shall describe another complication, which is caused by the rotating panel design of the LFS.

Each month a sample of about 6500 households is randomly selected from the Dutch population. These households are interviewed five times at quarterly intervals. Each month, data are collected in five different panels and each month five independent general regression estimates are obtained to estimate the monthly unemployed labour force.

A major problem with this rotating panel is that systematic differences between the subsequent panels occur. This is a well known problem for rotating panels and the literature refers to it as rotation group bias, Bailar (1975). The solid line in Figure 3 shows the general regression estimates for the monthly unemployed labour force at the national level based on the data collected in the first panel only. The dashed line is the means over the four general regression estimates of panel two through five. This illustrates that the unemployed labour force in the first panel is systematically larger compared to the four subsequent panels.

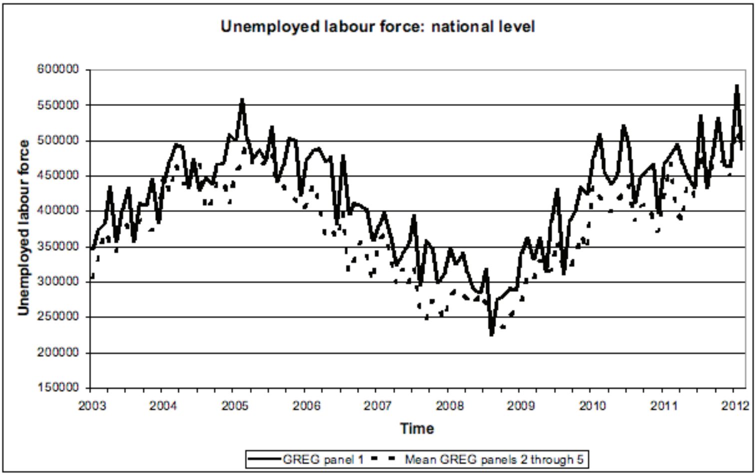


Figure 3: Rotation group bias unemployed labour force at the national level; comparison of the general regression (GREG) estimates based on the first panel (solid line) with the mean of the general regression estimates based on the four subsequent panels (dashed line).

The differences are the result of non-sampling errors like measurement errors and panel attrition. The time series model, already introduced in the example of small area estimation, uses the five series of the general regression estimates as input. The model accounts for the rotation group bias by benchmarking the domain estimates to the level of the series observed with the first panel, Van den Brakel and Krieg (2009).

In 2010 the LFS was redesigned. Briefly, the data collection in the first panel changed from face-tot-face interviewing to a mix of data collection modes that is based on telephone and face-to-face interviewing. Also the questionnaire in the different panels was adapted to the new data collection approach.

To test the effect on the main variables, the first panel of the old and the new design was conducted in parallel for a period of six months on a full scale. This test showed that the introduction of a new data collection mode, and a new questionnaire, increases the unemployed labour force at national level by 55,000 people. Discontinuities in the other panels are estimated by adding an intervention variable for each panel that models the moment that the survey process changed from the old to the new design and have values varying between 55,000 and 75,000.

Changes in field work methods and questionnaire design generally result in this kind of jump in the outcomes of the survey. This illustrates the sensitivity of the survey outcomes to measurement errors and the necessity to quantify these effects. If we did not quantify this effect of the redesign, then the figures about the unemployed labour force would wrongfully indicate an increase of 12%.

As a result, a time series model is obtained that uses the series with general regression estimates observed with the five panels as input. These series are plotted in Figure 4. It illustrates how noisy these five series are. This model, which is currently used for official publication purposes, accounts for:

- Small sample sizes by taking advantage of sample information observed in preceding periods.
- Rotation group bias by benchmarking the estimates to the level of the first panel.
- The discontinuities due to the redesign in 2010.

The model filters a signal, which is in this application defined as a trend plus a seasonal component, and a trend for the unemployed labour force from the five series of the general regression estimates, see Figure 4. Details are given by Van den Brakel and Krieg (2009, 2012).

Until 2010, the level of the filtered signal and the trend were equal to the level of the general regression estimates of the first panel, since the model removes the rotation group bias by benchmarking the outcomes to the level of the series obtained in the first panel. In 2010 the change-over to the new design started. As explained, the discontinuities resulted in higher levels for the series of general regression estimates of the five panels. In this application, the time series model estimates figures that is corrected for this discontinuity. As a result, the filtered signal and trend drops below the level of the series observed with the first panel after 2010.

The use of this model-based procedure in the production of official statistics is novel. Among other national statistical institutes, Statistics Netherlands is very innovative at this point.

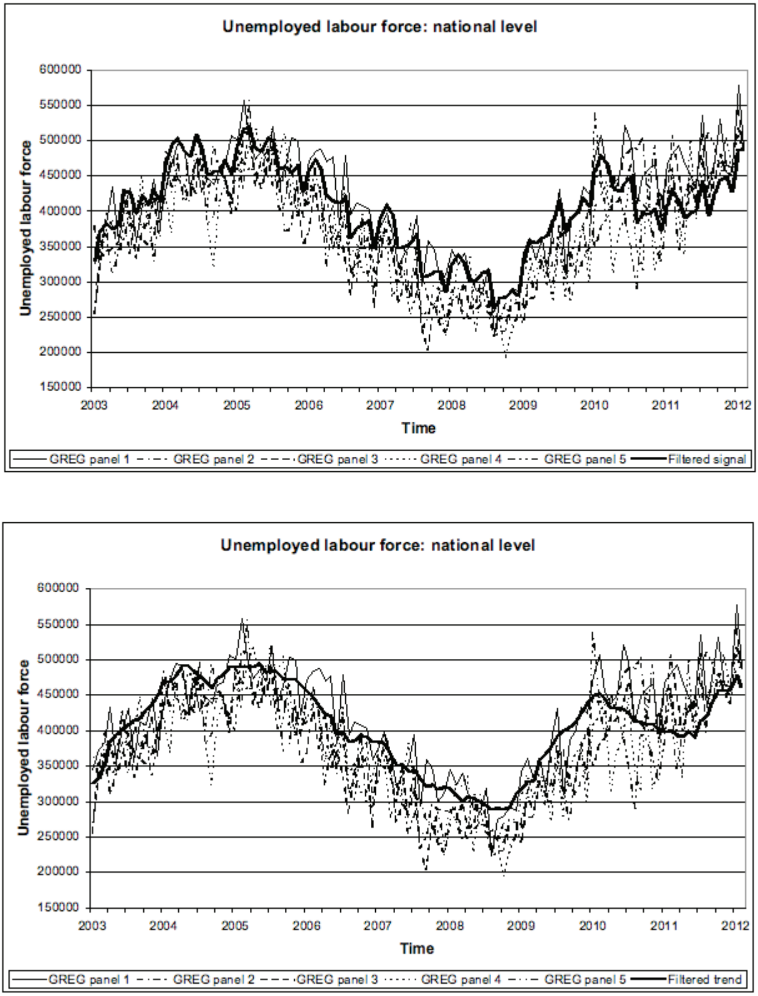


Figure 4: Unemployed labour force at the national level; general regression (GREG) estimates of the five panels and filtered signal (top panel) and trend (bottom panel) based on a structural time series model.

Alternative data sources for official statistics

There is a persistent pressure on national statistical institutes to reduce administration costs and response burden for businesses. This must be accomplished by using register data like tax registers, or other large

data sets that are generated as a by-product of processes unrelated to statistical production purposes. Examples are data available from mobile phone companies and social media like Twitter.
The process that generates the data might be selective with respect to the intended target population. One challenging problem in this context is to produce official statistics that are representative of the target population. There is no randomized sampling design that facilitates the generalization of conclusions and results obtained with the available data to an intended target population. As a result, the traditional design-based inference framework is not appropriate to these situations.

Model-based inference can be used if auxiliary information is available that explains the selectivity of the data. In many situations, however, the available auxiliary information will be limited. Buelens et al. (2012) explored the use of algorithmic inference procedures, like neural networks and regression and classification trees in a simulation study and concluded that such approaches might be beneficial if strong auxiliary information is lacking. Nevertheless, there is still a long way to go before this kind of data with these kind of methods can be used to produce official statistics.

Concluding remarks

In this talk I have emphasised the role of model-based inference in official statistics. The three applications that I have discussed do not constitute an exhaustive list. They are selected because they play an important role at national statistical institutes. The relation between these topics is that they share the same type of potential solutions, namely a more explicit use of statistical modelling.

The example of the monthly Labour Force Survey data illustrates how effectively econometric modelling of sample survey data simultaneously solves problems of small sample sizes, rotation group bias, and discontinuities. Alternative data sources might constitute an important alternative for survey data and its use in official statistics cannot be denied. They will, on the other hand, never completely substitute the traditional sample surveys since they simply do not supply the information about the wide range of topics about our society for which official statistics are required.

It can therefore be expected that the traditional design-based and model-assisted modes of inference will always play an important role in official statistics and that the importance of model-based inference and probably also algorithmic inference will rapidly increase.

The question remains how the reserved attitude with respect to the use of model-based procedures in official statistics can be changed into an attitude where this methodology is embraced. The hesitation to apply these methods probably finds its origin in the unfamiliarity with this methodology and the fear that these methods are less robust for model-misspecification.

One way to overcome these obstacles is to increase our knowledge and experience of the application of these methods in the context of official statistics. This is one of the reasons that Statistics Netherlands and Maastricht University have installed this chair in Survey Methodology. It increases the interface between academic research and the more practical world of official statistics. Hopefully this will stimulate the application of advanced econometric methods in this context.

Acknowledgement

We have come to the end of my speech. Personally I want to thank Statistics Netherlands for offering me this challenging scientific career opportunity. I would like to thank the Rector Magnificus, the members of the board of the School of Business and Economics, the Dean Jos Lemmink, the head of the Econometrics group Franz Palm, the Director General of Statistics Netherlands Gosse van der Veen, Kees Zeelenberg, Wim de Witte, Frank van de Pol, and all the others who facilitated my appointment, and for the confidence they place in me.

I am looking forward to an interesting and challenging period where I can cooperate with my colleagues from Statistics Netherlands as well as Maastricht University in many inspiring projects.

I would like to remember my father and mother who did everything within their power to motivate me and support me in reaching my goals. I regret that they cannot join us on this beautiful day.

Finally I would like to thank Marion for supporting me in my ambitions and my two sons Luc and Sven who are always a great source of pleasure. Particularly during difficult or stressful periods, a smile on the face of my sons puts everything into the right perspective.

Ladies and gentlemen, thank you for your attention. I have spoken.

References

Alexander, C.H. (1987). A class of methods for using person controls in household weighting. *Survey Methodology*, 13, pp. 183-198.

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.

Battese, G.E., R.M. Harter, and W.A. Fuller (1988). An error components model for prediction of county crop areas using satellite data. *Journal of the American Statistical Association*, 83, pp. 28-36.

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

Boonstra, H.J., J.A. van den Brakel, B. Buelens, S. Krieg, and M. Smeets (2008). Towards small area estimation at Statistics Netherlands. *Metron International Journal of Statistics*, LXVI, pp. 21-50.

Boonstra, H.J., B. Buelens, K. Leufkens, and M. Smeets (2010). Small area estimates of labour status in Dutch municipalities. Research paper, BPA nr. DMH-2010-10-29-hbta, Statistics Netherlands, Heerlen.

Buelens, B., H.J. Boonstra, J.A. van den Brakel, and P. Daas (2012). Shifting paradigms in official statistics; from design-based to model-based to algorithmic inference. Working paper, Statistics Netherlands, Heerlen.

Buelens, B., and J.A. van den Brakel (2011). Inference in surveys with sequential mixed-mode data collection. Discussion paper 201121, Statistics Netherlands, Heerlen.

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

Fay, R.E., and R.A. Herriot (1979). Estimation of income for small places: an application of James-Stein procedures to census data. *Journal of the American Statistical Society*, 74, pp. 268-277.

Hansen, M.H., and W.N. Hurwitz (1943). On the theory of sampling from finite populations. *Annals of Mathematical Statistics*.14, pp. 333-362.

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

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.

Kempthorne, O. (1955). The randomization theory of experimental inference. *Journal of the American Statistical Association*, 50, pp. 946-967.

Krieg, S., and J.A. van den Brakel (2012). Estimation of the monthly unemployment rate for six domains through structural time series modelling with cointegrated trends. *Computational Statistics and Data Analysis*, in press.

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

Neyman, J. (1934). On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society*, 97, pp. 558-625.

Pfeffermann, D. (1991). Estimation and seasonal adjustment of population means using data from repeated surveys. *Journal of Business & Economic Statistics*, 9, pp. 163-175.

Pfefferman, D. (2011). Modelling of complex survey data: Why model? Why is it a problem? How can we approach it? *Survey Methodology,* 37, pp. 115-136.

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.

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

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.

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, Series A*, 171, pp. 581-613.

Van den Brakel, J.A., B. Buelens, and H.J. Boonstra (2012). Small area estimation to quantify discontinuities in sample surveys. Research paper, BPA nr.: PPM-13-04-2012-JBRL, Statistics Netherlands, Heerlen.

Van den Brakel, J.A., and S. Krieg (2009). Estimation of the monthly unemployment rate through structural time series modelling in a rotating panel design. *Survey Methodology*, vol. 35, pp. 177-190.

Van den Brakel, J.A., and S. Krieg (2011). Estimation of monthly unemployment figures in a rotating panel design; on the use of auxiliary series in structural time series models. Discussion paper 201122, Statistics Netherlands, Heerlen.

Van den Brakel, J.A., and S. Krieg (2012). Dealing with small sample sizes, rotation group bias and discontinuities in a rotating panel design. Working paper, Statistics Netherlands, Heerlen.

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

Van den Brakel, J.A., and J. Roels (2010). Intervention analysis with state-space models to estimate discontinuities due to a survey redesign. *Annals of Applied Statistics*, 4, pp. 1105-1138.

Ybarra, L.M.R., and S.L. Lohr (2008). Small area estimation when auxiliary information is measured with error. *Biometrika*, 95, pp. 919-931.

OBS 8605 Design en print Océ Business Services, Maastricht