Catalogue no. 12-001-X
ISSN 1492-0921

Survey Methodology

Comments on “Jean-Claude Deville’s contributions to survey theory and official statistics”

by Carl-Erik Särndal

Release date: January 3, 2024

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at infostats@statcan.gc.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

• Statistical Information Service	1-800-263-1136
• National telecommunications device for the hearing impaired	1-800-363-7629
• Fax line	1-514-283-9350

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

An HTML version is also available.

Cette publication est aussi disponible en français.

Survey Methodology, December 2023
Vol. 49, No. 2, pp. 339-345
Statistics Canada, Catalogue No. 12-001-X

Comments on "Jean-Claude Deville's contributions to survey theory and official statistics"

Carl-Erik Särndal[1]

Abstract

In recent decades, many different uses of auxiliary information have enriched survey sampling theory and practice. Jean-Claude Deville contributed significantly to this progress. My comments trace some of the steps on the way to one important theory for the use of auxiliary information: Estimation by calibration.

Key Words: Auxiliary information; Generalized regression estimation; Calibration estimation.

Introduction

I am honored to reflect on the article by Ardilly, Haziza, Lavallée and Tillé on Jean-Claude Deville and his many contributions to survey theory and practice.

I refer to the article as Ardilly, Haziza, Lavallée and Tillé (AHLT). One part of it is devoted to estimation of the finite population total by calibration theory. I concentrate on it here. More than thirty years have passed since the publication of the influential article on calibration by Deville and Särndal (1992). It was followed a year later by an important sequel, also published in the Journal of the American Statistical Association (JASA).

A forerunner – and a special case – of the calibration estimator is "the well-known estimator by generalized linear regression", as AHLT phrase it. Given in their formula (3.5), it is popularly known as the (linear) generalized regression (GREG) estimator. This gives me reason here to examine how the GREG estimator paved the way for the calibration estimator. The GREG estimator evolved in form and shape in the mid-to-late seventies and in the eighties; it is fair to say that it is the progenitor of the estimator by calibration.

The GREG construction was a product of "the prediction argument": To predict as well as possible the study variable values for the non-observed population units. On the other hand, a somewhat later stream in the literature dwells on "the weighting argument": To conceive and compute the estimator of the population total through an appropriate weighting of the study variable values y_k observed for units k in the probability sample s, preferably a weighting more "information laden" than simply the inverse inclusion probability weighting $d_k = 1/\pi_k$.

My comments here trace my gradual understanding over several decades of estimation in the presence of auxiliary information. They are an incomplete testimony to a period of development where Jean-Claude Deville played an important part. Many others contributed and should have been recognized in my notes.

1. Carl-Erik Särndal, professor emeritus, Statistics Sweden, Lilla Norregatan 12, SE 27135 Ystad, Sweden. E-mail: carl.sarndal@telia.com.

Meeting

I first met Jean-Claude in the Swiss Alps. Not on skis, but at the traditional yearly "winter school" in statistics held at the resort Les Diablerets. We were both invited speakers there in 1987. I had been told beforehand that I was going to meet a French statistician, relatively new to the survey sampling field, who had made himself known by some recent interesting contributions.

I did meet Jean-Claude there, in person, and in spirit. Contact and mutual understanding happened from the beginning. As a graduate of the prestigious Ecole Polytechnique, Jean-Claude relied on a strong mathematical background and ability; in addition, his work at Institut national de la statistique et des études économiques (INSEE) made him familiar with national statistical agencies and their efforts to produce accurate national statistics.

A unique combination seemed to me to characterize him, as I got to know him: There was a mathematical insight that he applied in sometimes astonishing ways, yet with a clear background in national statistics production. He was very sensitive to resolving problems embedded in a greater practical environment, such as a national statistical agency.

As he explained to me, at already well over 30, he had decided to change fields of research interest to survey science. The idea of probability sampling intrigued him: to select, with known probabilities, from a finite set of identifiable units. His curiosity was stimulated by the amazing variety of ways to do that, and how to estimate the population parameters on data from a probability sample using auxiliary variables.

Seven years his senior, I had behind me a longer exposure to the field. He told me how he had systematically "taken in" the field of survey science, had studied the significant material, from the classical sampling books of the fifties to the recent work, in the newborn vigor that survey sampling theory found from around 1970. For example, he admired the work of the Czech sampling theoretician Jaroslav Hàjek. He had studied the 1977 Wiley book that I had co-authored, *Foundations of Inference in Survey Sampling.* In the years that followed, he came to Canada several times and we worked together at the Université de Montréal.

Randomization theory

I believe that Jean-Claude's ideological preference agreed, essentially, with the randomization theory, the design-based fold of survey science. That is, an approach where the inference about the finite population is built on the probability sampling design, using the known inclusion probabilities of the sampled units.

This positioning was by no means obvious, for someone who was learning the field in those days. The 1970's had brought a flux of ideas, some of them conflicting. A formidable challenger to the traditional design-based randomization theory was the model-based theory, which maintained instead that inference was to be based on the stochastic structure stipulated in the assumed model.

So it seems to me that when Jean-Claude set out to learn the field, he might well have felt attracted to, and become a convinced disciple of, the model-based camp. But he did not.

Auxiliary information

Estimation built around ideas of supplementary information and predicted y-values for non-observed units seems to have been studied, or attempted, as early as in the 1950's, in places, such as the national statistical agencies, where insightful survey methodologists dwelled. At INSEE in France, P. Thionet and Y. Lemel were among persons whose important work influenced Jean-Claude.

An important instance of auxiliary information occurs when values are known for all population units on one or more variables, x-variables, thought to be related to the survey variable, the y-variable. The Scandinavian countries, equipped as they are with high quality population registers of different kinds, were testing grounds for this methodology. The uses of auxiliary information led, beginning in the 1970's, to a strong development in survey statistics research and practice.

Generalized regression estimator

The GREG estimator rests on "the prediction argument"; the y-values for the non-sampled units are predicted, on the basis of a perceived relationship between the study variable y and one or more correlated auxiliary variables. The linear GREG estimator, as we know it today, given in formula (3.5) in AHLT, evolved gradually, from the mid-to-late 1970's onward. Several took a part in this; a gradual refinement took place. The first time that I used "generalized regression estimator" as a term and as a construction principle – incomplete at the time – was, as far as I recall now, in a co-authored article in Biometrika, Cassel, Särndal and Wretman (1976). Among contributions to "the regression thinking", the work at Iowa State by W.A. Fuller and his students stands out.

"Generalized" meant essentially that the early form of the GREG estimator extended what Cochran and other texts from the 1950's had presented as a method to reduce variance, by attaching a linear regression adjustment term to the basic design-unbiased estimator. The adjustment, small in magnitude in large samples, was a function of the auxiliary variables.

More generally, let $\hat{y}_k$ be the predicted value, by a linear or non-linear assisting model, of the study variable value y_k. The GREG construction is $\hat{Y}_{\text{GREG}} = \sum_s d_k y_k + \left(\sum_U \hat{y}_k - \sum_s d_k \hat{y}_k \right)$, with U representing the population and s representing the probability sample from U. The residuals $y_k - \hat{y}_k$ left by the model fit become more apparent when we write it as $\hat{Y}_{\text{GREG}} = \sum_U \hat{y}_k + \sum_s d_k (y_k - \hat{y}_k)$.

In papers from the late seventies and early eighties, I had occasion to examine the linear GREG estimator, that is, where $\hat{y}_k$ is the result of a linear regression fit. I remember the 1976 co-authored Biometrika article especially because of an administrative curiosity.

Although it was exceptional for an ordinary submission to Biometrika, the chief editor, D.R. Cox, wished to accompany the article with a special discussion, something to which we agreed. The article was seen, apparently, as an unorthodox mixture of an appeal to a model and its properties, and, at the same time, to the randomization distribution, p, arising from the probability sampling.

The discussant was T.M.F. (Fred) Smith, a highly respected survey theoretician with whom I came to enjoy, as time went by, much fruitful and friendly contact. He wrote: “This paper raises a fundamental issue in finite population inference via a superpopulation. … The authors impose the further constraint that (the regression estimator) T should be p-unbiased … Why should the selection probabilities, p, take any precedence over the model ξ?” It was a well-motivated question from his model-based point of view: If the regression model was trustworthy for building the estimator, why should it then be abandoned, in favour of the randomization distribution, when it came to evaluating basic properties, such as bias and variance?

This was an illustration of an exchange that could occur in those days, between a member of the new model-based camp and one of the traditional design-based camp. In due time, the perspective in that article developed into “model-assisted design-based estimation”, i.e., assisted by the model but not dependent on its “truth”. The asymptotic p-unbiasedness served as protection against a possibly false or improper model. The “truth” or not of the model was not the primary issue. If the model does not hold, the estimator is still asymptotically design-consistent.

Reoriented thought process

The literature from 1980 and later suggested a “reorientation of the thought process”; some of the attention shifted away from the prediction argument to an alternative argument centered on the weights assigned to the observed sample y-values.

In the prediction approach, I and others had built an estimator by predicting the unobserved y-values as well as possible, via a regression fit of some kind, and with a use of the auxiliary variables. In the interest of accurate estimation, it was clear that the predictions $\hat{y}_k$ must reflect as accurately as possible the unknown y_k for the non-observed units; the residuals $\hat{y}_k - y_k$ should be small.

The weighting argument focuses instead on estimation by an appropriate weighting of the y-values observed in the sample. In the interest of accurate estimation, the weight given to a sampled unit should reflect what is known, and what is particular, about a unit in the population. “Good weights” could or should be sample-dependent functions of the auxiliary vector values $\mathbf{x}_k$. They might entail just a small but important adjustment to the basic design weights $d_k = 1/\pi_k$, as is the case with the weights implied by the GREG estimator formula.

In Särndal (1982), I write the linear GREG estimator as a weighted sum over the sample, with the observed study variable value y_k receiving the weight $d_k g_k$, where the sampling design weight $d_k = 1/\pi_k$ undergoes a small adjustment g_k, slightly away from “one”. Bethlehem and Keller (1987) also show the linear weighting interpretation of the GREG estimator, and they note the calibration property of the weights.

Model-assisted survey sampling

The thick manuscript of the book with that title, by Särndal, Swensson and Wretman, was nearing completion when I and Jean-Claude co-operated in the late 1980’s. I described to him the spirit of the book.

Published in 1992 by Springer-Verlag; it became popularly known as the Yellow Book. It advocated a design-based outlook on inference, more particularly in a vein that became widely known as model-assisted. The role of the model was explained at length, in particular in Chapter 6; the wording was important; it was convincing and inspired others.

Curious as it may seem today, as book authors we were quite sensitive to the survey theory climate. For or against "reliance on models" had been a much debated question in the survey sampling theory literature since the early 1970's. We had had our part of "the hot feelings", through the case with the above mentioned 1976 Biometrika article on generalized regression estimation.

The Yellow Book used the prediction argument to carefully build and explain the GREG estimator of $Y = \sum_U y_k$. Predicted values $\hat{y}_k$ from a linear regression fit are given by $\hat{y}_{ks} = \mathbf{x}_k' \mathbf{B}_s = \mathbf{x}_k' \left(\sum_s d_k \mathbf{x}_k \mathbf{x}_k'\right)^{-1} \left(\sum_s d_k \mathbf{x}_k y_k\right)$, where s is the probability sample and $d_k = 1/\pi_k$. They are computable for all units if $\mathbf{x}_k$ is available for all. The linearly weighted form of the GREG estimator is also emphasized:

$$\hat{Y}_{\text{GREG}} = \sum_s d_k g_k y_k,$$

where $g_k = 1 + \left(\sum_U \mathbf{x}_k - \sum_s d_k \mathbf{x}_k\right)' \left(\sum_s d_k \mathbf{x}_k \mathbf{x}_k'\right)^{-1} \mathbf{x}_k$.

The gee-weights

The "gee-weight", as we used to call g_k, is a weight factor, equaling one plus a term of minor magnitude in large samples, nevertheless with an important impact. It modifies slightly the design weight $d_k = 1/\pi_k$ into a total weight $d_k g_k$, and, as the Yellow Book explains, those weights satisfy

$$\sum_s d_k g_k \mathbf{x}_k = \sum_U \mathbf{x}_k.$$

It came to be known as the calibration property. I vividly remember my momentary surprise at my own "discovery" of the property; it was, however, quite evident upon a closer look, and therefore not given any immediate attention. Others active in the field were no doubt also aware of it in the years around 1980. The weights $d_k g_k$, their function, their calibration property, and their use in variance estimation were at the center of attention in a Biometrika article by Särndal, Swensson and Wretman (1989).

The calibration property of the weights $d_k g_k$ was, however, a signal for my interest in calibrated weight systems: it must be a more generally fruitful idea. When Jean-Claude and I discussed it in the late 1980's, the property was a well-established fact, a starting point: It must be possible to extend and generalize it.

The rise of calibration theory

A phrase in AHLT catches my attention: The authors maintain that the calibration theory, as in Deville and Särndal (1992), brought "… one of the most important advances in the field of estimation in the presence of auxiliary information: It is possible to construct the GREG estimator by means of a calibration." In other words, a "discovery" was that calibration theory is sufficiently broad in scope to admit the important GREG estimator under its umbrella.

It is indeed important to have the GREG estimator as a member of the calibration family; I add, though, that my own insight happened in a different temporal order: Knowing in advance that the GREG weights have the calibration property, it must be possible to extend the idea.

This took form in the theory well described in AHLT, with a measure of distance to minimize, between the survey weights $d_k = 1/\pi_k$ and new weights w_k, subject to the calibration constraint $\sum_s w_k \mathbf{x}_k = \sum_U \mathbf{x}_k$. A resulting calibrated weight w_k is a function of the auxiliary vector value $\mathbf{x}_k$.

The calibration theory: Implications and interpretations

I started these comments by noting two ways to proceed, "the prediction argument" as opposed to "the weighting argument". This important distinction has set its mark on a part of theory development in survey sampling over the past several decades.

The former was successfully used in creating the (linear or non-linear) GREG estimator. It seems to me now that the weighting argument, as used in calibration theory, has a broader application, or a wider appeal. To estimate, you just add up the properly weighted y-values.

Much has been said on calibration in the literature since the Deville and Särndal (1992) article. It is true that calibration was presented there as a design-based methodology under ideal survey conditions, one hundred percent survey response, and an absence of other survey errors as well. However, time passing, the concept of calibration has shown itself to be an instrument of extraordinary power and flexibility, as witnessed, for example, in the variety known as model calibration, and especially in its extensions to nonresponse adjustment, treated in a large literature of its own.

An image

I close with a digression on the word "calibration". At INSEE, the French term "calage" was apparently well established long ago. It was used, it seems, in early examples of a weighting that confirms the known population total of an auxiliary variable.

As a French verb, "caler" means to rig, to fix, to stabilize. Elsewhere also, there were no doubt insightful statisticians who derived and computed weights, with the property today called "calibrated", well before there was a special name and a special theory for the procedure.

For the 1992 JASA article, Jean-Claude and I settled on the term "calibration", aware of the meaning it has to some, namely, in the idea of the balance scale, this instrument used in food stores long ago: two plates at either end of beam. To fill a customer's order for, say, 600 grams of coffee or butter or flour, the store clerk placed weights on one plate, amounting to 600 grams, then measured up the desired food item on the opposite plate, until balance occurred. The store had a collection of metal weights, in different denominations; they were certified, calibrated, to guarantee the customer's right to a correct weighting. As some have fondly reminded me over the years, “calibrated weights” brings up this old mental image of a trustworthy procedure.

References

Bethlehem, J.G., and Keller, W.J. (1987). Linear weighting of sample survey data. *Journal of Official Statistics*, 3, 141-153.

Cassel, C.M., Särndal, C.-E. and Wretman, J.H. (1976). Some results on generalized difference estimation and generalized regression estimation for finite populations. *Biometrika*, 63, 615-620.

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

Särndal, C.-E. (1982). Implications of survey design for generalized regression estimation of linear functions. *Journal of Statistical Planning and Inference*, 7, 155-170.

Särndal, C.-E., Swensson, B. and Wretman J.H. (1989). The weighted residual technique for estimating the variance of the finite population total. *Biometrika*, 76, 527-537.

Särndal, C.-E., Swensson, B. and Wretman J.H. (1992). Model assisted survey sampling. New York: Springer-Verlag.