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

Survey Methodology

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

by Françoise Dupont

Release date: January 3, 2024

Statistics Canada Statistique Canada

Canada

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

© His Majesty the King in Right of Canada, as represented by the Minister of Industry, 2024

All rights reserved. Use of this publication is governed by the Statistics Canada Open Licence Agreement.

An HTML version is also available.

Cette publication est aussi disponible en français.

Survey Methodology, December 2023
Vol. 49, No. 2, pp. 313-319
Statistics Canada, Catalogue No. 12-001-X

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

Françoise Dupont[1]

Abstract

Many things have been written about Jean-Claude Deville in tributes from the statistical community (see Tillé, 2022a; Tillé, 2022b; Christine, 2022; Ardilly, 2022; and Matei, 2022) and from the *École nationale de la statistique et de l'administration économique* (ENSAE) and the *Société française de statistique*. Pascal Ardilly, David Haziza, Pierre Lavallée and Yves Tillé provide an in-depth look at Jean-Claude Deville's contributions to survey theory. To pay tribute to him, I would like to discuss Jean-Claude Deville's contribution to the more day-to-day application of methodology for all the statisticians at the *Institut national de la statistique et des études économiques* (INSEE) and at the public statistics service. To do this, I will use my work experience, and particularly the four years (1992 to 1996) I spent working with him in the Statistical Methods Unit and the discussions we had thereafter, especially in the 2000s on the rolling census.

Key Words: Methodology; Calibration; Balanced sampling; Automatic coding; Harmonic analysis; Jean-Claude Deville.

I first learned of Jean-Claude Deville's work on harmonic analysis of data on the trajectories of women's fertility for my studies at ENSAE in the late 1980s. Jean-Claude Deville had examined these methods in the 1970s.

In his 1974 article (Deville, 1974), Jean-Claude Deville explained that after examining fertility in previous studies, he was looking for a satisfactory exploratory method to use rich individual data since many were available for multiple dates, which was quite rare at the time. He used the 1962 family survey, which provided information on the life trajectory of female respondents regarding the birth of children. In his introduction "Statistician seeking an analytical method," he explains that he wants to analyze the gradual formation of families and the calendar of births based on the number of years married and that his entire thinking stems from a concrete problem with the data to analyze. He explains in detail why he does not want to use econometric methods that presuppose a model – and therefore a preconceived idea of causal links. We already see his reservations about econometrics, which increased later on. Next were the mathematical developments to show how to adapt the general principle of harmonic analysis to time series data. This is an extension of harmonic analysis, as it is usually practised. An application of this was illustrated in the article by Deville (1977).

This is a typical example of the research process in applied statistics: we start with a tangible case in a work environment for which we do not yet have a suitable tool, then conduct bibliographical research by integrating our predecessors' contributions, and lastly, a mathematical reflection to propose a new method to test out in real life, refine and popularize. This long, open process, the example of Pierre Thionet's career (1916-2002, architect of survey methodology at INSEE from 1946, to whom we owe the introduction of random surveys at INSEE, together with Raymond Lévy-Bruhl, who came up with methodological

1. Françoise Dupont, Demographic and Social Statistics Directorate, Résil Program, INSEE, 88 avenue Verdier, 92120 Montrouge. E-mail: francoise.dupont@insee.fr or frdupont-dom@orange.fr.

reflections involving many ideas that went on to be developed) and whom Jean-Claude admired (Deville, 2003; Armatte, 2003; Ardilly, 2022), explains why he has always defended career-based logic for methodologists so that ideas can come from accumulated experience and feed on theoretical readings, by proposing improved methods in order to innovate.

For him, the elegance of the mathematical solution and its successful application were a source of great satisfaction. He spoke passionately about the findings he was proud of, such as drawing 1 out of 20 people for the 1990 census, which already comprises the idea of balanced sampling that Tillé (2022a) spoke about in his tribute to Jean-Claude Deville at the 2022 *Journées de Méthodologie Statistique* conference (Statistical Methodology Days). Another source of pride was the survey designs for external requesters for which he proposed creative, custom-made solutions.

He was the architect of two "institutional" creations that expanded over time: a Statistical Methodology Unit and the *Journées de Méthodologie Statistique*. In both cases, "problems and solutions" were prioritized over "theory."

Creation of the Statistical Methodology Unit

In 1994, the Statistical Methodology and Survey Division saw its missions expanded and its staff grow, going on to become the *Unité Méthodes Statistiques* (UMS) (Statistical Methodology Unit), which was responsible for the statistical methods used to produce demographic and social statistics. He led the UMS until 1998.

Journées de Méthodologie Statistique

Jean-Claude Deville created the *Journées de Méthodologie Statistique*, based on the concept of the annual scientific meetings that take place at the U.S. Census Bureau and Statistics Canada. In his view, there were a number of objectives, which he presented in his introduction of the third edition in 1993 (Deville, 1993):

"*There are many general statisticians in regional directorates and departmental statistical services with few opportunities to meet to talk shop.*" [translation]. The purpose was to lead a network of statisticians and maintain their skills.

"*The second, harder idea […] was to give visibility to methodology work.*" [translation]

"*Lastly, the third idea: ask foreign colleagues to give their thoughts on a specific problem and comment on INSEE's work.*" [translation]

He concludes as follows:

"*In the end, the success of the Journées is easily explained: official statistics, as a discipline, exist in both France and around the world. It requires effort; the Journées de la statistique officielle are the periodic scientific meeting. It is easy to think of what is missing: an associated structure, as well as training and research structures in universities. All this should result in a more systematic publication of the work in more accessible forms than the current ones.*" [translation]

Of the four years (1992 to 1996) I spent in the Statistical Methods and Survey Division and in the UMS, in two different positions, I witnessed many changes to statistical practices, always based on two pillars: theory and practice.

In 1992, when I returned to the UMS to work on non-response, the theoretical groundwork of calibration was established and the SAS CALMAR macro (Deville, Särndal and Sautory, 1993) developed by Olivier Sautory could already be used to apply this method. On one hand, the tangible work of INSEE statisticians consisted in applying an iterative calibration procedure by successively calibrating each margin until the process converges to adjust their surveys.

Thanks to the first article on calibration published in 1992 in JASA (Deville and Särndal, 1992), he scored very high points with his INSEE colleagues and gave survey methodology its first distinctions at the institute. It had yet to be proven that this method simultaneously corrected non-response bias and improved the accuracy of the estimator, as per the intuition of practitioners and their survey adjustment practices. This is the work I completed with Jean-Claude Deville, which resulted in publications for the 1993 *Journées de méthodologie statistique* (Deville and Dupont, 1993; and Dupont, 1993).

We also needed to expand the use of the SAS CALMAR macro among practitioners. This is why the UMS gave many presentations on the CALMAR macro and offered to advise survey managers. Thanks to the presentations, the availability of the tools and accompanying advice, and management of the adjustment of complex surveys, we were able to gain the trust of survey managers and gradually modify their practices. This led to more in-depth reflection on the different cases of adjustment and to moving toward more homogeneous practices and a more sophisticated solution, when applicable (Dupont, 1995).

In the same spirit, other practices were developed with the help of his UMS colleagues. The growing success of the methodology days demonstrated the desire and need to talk about and share statistical methodology in all fields, without prioritizing subjects according to their mathematical content, but rather according to their actual usefulness, in a field wider than econometrics, the only statistical discipline truly recognized at INSEE at the time. He helped to bolster the UMS's credibility and advance its recognition in official statistics in France and around the world.

This recognition was confirmed with the development of the unit in charge of demographic and social statistical methodology, whose human and financial resources expanded in 1994, while opening up to data collection methods with the creation of a small team to explore this, which he entrusted to me.

Under his leadership, the UMS conducted purely methodological operations along with current collection operations. Work was initiated to quantify the interviewer effect (Berthier, Deville and Néros, 1998) and two other projects to analyze the interviewer effect more qualitatively in surveys being developed on the duration of work and situations of disadvantage, another on the effect of mandatory response (Berthier and Dupont, 1999), and lastly, a more thorough project measured the impact of manual editing by managers on paper questionnaires. This involved examining the reality of the manual edits done in regional directorates based on instructions from survey designers in order to examine the evolution of edits with a view to transitioning surveys to computer-assisted personal interview (CAPI) mode, which includes the

possibility of automatic edits. By looking at the differences between the instructions given and the actual work done, this project illustrated the difficulty involved in manual edits.

Many other projects began during this period in order to advance practices through turnkey tools provided to statisticians, as is already done at Statistics Canada.

After the CALMAR tool, management of the automatic coding tool SICORE, designed by Pascal Rivière and Eric Meyer (Rivière, 1993) and led by Jean-Claude Deville, was transferred to the UMS to develop and systematize its use for coding occupations and countries. The tool was also adopted by foreign countries through INSEE's cooperation activities.

In 1994, a *Comité du label* (approval committee) was created (Christine and Roth, 2020), with the purpose of looking at survey design in terms of sampling, adjustment, questionnaire design and, more generally, the information collection system from a methodology perspective. The UMS was represented by an expert in charge of reviewing how best practices in terms of methods are applied. This systematic review of the methodology of all surveys helped to gradually develop the role of advising survey designers on the adjustment and estimation phases, as well as on questionnaire design (Bilocq, 1996). He contributed to advancing tangible practices and developing the advisory role of the UMS in survey development, which was formalized a few years later in the organization.

The UMS took on the management of a common set of questions in household surveys about household composition, which had just been developed by a joint effort led by an experienced survey designer. The challenge was to systematize its use by offering services, as for a tool, in order to harmonize practices inasmuch as possible. Variables were added to produce data on immigrants and foreigners that could be analyzed in all household surveys. Managing the conversion to CAPI and automatic coding has made it more attractive to survey designers, and has enabled us to take pooling in surveys further, thanks to a turnkey tool.

In line with the idea of a turnkey tool and advice on using it, the design of a precision calculation software package called POULPE (Caron, Deville and Sautory, 1998) was launched.

In the same period, the methodology of the employment survey was undergoing a major change, with the introduction of an ongoing employment survey. The UMS played a key role in the discussions on the various aspects of methods and organization (Lagarenne and Schuhl, 1995; Détour, Thiesset and Schuhl, 1995) that led to the introduction of the ongoing employment survey, which is still conducted today.

The need for local statistics incited the UMS to support the implementation of a local population estimation method (Decaudin and Labat, 1996).

At the same time in 1994, Jean-Claude Deville and other INSEE colleagues were thinking about making methodological changes to the census, which was to later become the "rolling census" (Deville and Jacod, 1996). This unique method, though complex with regard to estimation and variance calculation, had many benefits for the organization, particularly smoothing the budget and the workload, which is why it was adopted (Durr, 2005). This transformation of the census was the first opportunity to use balanced sampling

when it was introduced in the 2000s, thanks to the CUBE macro (Tillé, 2022a). In practice, it revealed other qualities; for example, it contributed to implementing a cycle of ongoing improvements to the processes, which was not part of its initial objectives.

Everyone knows Jean-Claude Deville differently, based on their experience working with him. I especially wanted to show his constant desire to bring about changes to the work at INSEE, be it for "hard" methodology (estimation, adjustment, sampling and variance calculation) or "soft" methodology, i.e., collection methods that are harder to analyze, such as a topic of study, and are published much less often.

Thanks to his perseverance and the solutions he implemented, he helped survey methodology at INSEE gain more recognition. For Jean-Claude Deville, recognition of the role of methodology remained imperfect because the methodology used for businesses was the responsibility of another unit at the time, despite the potential for pooling resources.

It was his passion for theoretical mathematics and his unwavering belief in the day-to-day application that contributed to all the technical progress mentioned above and paved the way to an ambitious, recognized methodology at INSEE.

References

Ardilly, P. (2022). Hommage à Jean-Claude Deville. *Statistique et Société*, 10, 121-123, www.statistique-et-societe.fr.

Armatte, M. (2003). Pierre Thionet et l'introduction en France des méthodes de sondage aléatoire. *Journal de la société française de statistique*, 144, 227-255, http://eudml.org/doc/199757.

Berthier, C., Deville, J.-C. and Néros, B. (1998). Une méthode de mesure de l'effet enquêteur. *Actes des Journées de Méthodologie Statistique*, 133-143, Insee, Paris.

Berthier, C., and Dupont, F. (1999). L'incidence du caractère obligatoire des enquêtes. *Enquêtes et sondages : méthodes modèles et applications*, Dunod.

Bilocq, F. (1996). Conception et évaluation de questionnaire. *Actes des Journées de Méthodologie Statistique* 1996, Insee, Paris.

Caron, N., Deville, J.-C. and Sautory, O. (1998). Estimation de précision de données issues d'enquêtes : document méthodologique sur le logiciel POULPE. Methodological document on the POULPE software, Insee, Paris.

Christine, M. (2022). Discours introductif en hommage à Jean-Claude Deville. *Actes des Journées de Méthodologie Statistique*, Insee, Paris.

Christine, M., and Roth, N. (2020). Le Comité du label, Un acteur de la gouvernance au service de la qualité des statistiques publiques. *Courrier des statistiques*, 5, 39-52.

Decaudin, G., and Labat, J.-C. (1996). Une méthode synthétique, robuste et efficace pour réaliser des estimations locales de populations. *Actes des Journées de Méthodologie Statistique*, Insee, Paris.

Détour, C., Thiesset, C. and Schuhl, P. (1995). Contrôle de qualité de l'enquête trimestrielle emploi : résultats de l'enquête transitions sur le marché du travail. *Actes des Journées de Méthodologie Statistique*, Insee, Paris.

Deville, J.-C. (1974). Méthodes statistiques et numériques de l'analyse harmonique. *Annales de l'Insee*, 15, 3-101.

Deville, J.-C. (1977). Analyse harmonique du calendrier de constitution des familles en France, disparités sociales et évolution de 1920 à 1960. *Population*, 32, 17-63.

Deville, J.-C. (1993). Introduction. *Actes des Journées de Méthodologie Statistique*, Insee, Paris.

Deville, J.-C. (2003). In Memoriam Pierre Thionet 1915 – 2002. *The Survey Statistician*, 47, 2.

Deville, J.-C., and Dupont, F. (1993). Non-réponse : Principes et méthodes. *Actes des Journées de Méthodologie Statistique*, 53-69, Insee, Paris.

Deville, J.-C., and Jacod, M. (1996). Replacing the traditional French census by a large-scale continuous population survey. *Proceedings of the Annual Research Conference*, U.S. Census Bureau, Washington.

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

Deville, J.-C., Särndal, C.-E. and Sautory, O. (1993). Generalized raking procedure in survey sampling. *Journal of the American Statistical Association*, 88, 1013-1020.

Dupont, F. (1993). Calage et redressement de la non-réponse totale : Validité de la pratique courante de redressement et comparaison des méthodes alternatives pour l'enquête sur la consommation alimentaire de 1989. *Actes des Journées de Méthodologie Statistique*, Insee, Paris.

Dupont, F. (1995). Alternative adjustments where there are several levels of auxiliary information. *Survey Methodology*, 21, 2, 125-135. Paper available at https://www150.statcan.gc.ca/n1/en/pub/12-001-x/1995002/article/14397-eng.pdf.

Durr, J.-M. (2005). The French new rolling census. *Statistical Journal of the United Nations Economic Commission for Europe*, 22, 3-12, DOI:10.3233/SJU-2005-22102.

Lagarenne, C., and Schuhl, P. (1995). Contrôle de qualité de l'enquête trimestrielle emploi : résultats de l'enquête protocole. *Actes des Journées de Méthodologie Statistique*, Insee, Paris.

Matei, A. (2022). In memory of Jean-Claude Deville. *International Association of Survey Statisticians*, Webinar 17.

Rivière, P. (1993). SICORE : Système général de chiffrement automatique. *Actes des Journées de Méthodologie Statistique*, Insee, Paris.

Tillé, Y. (2022a). Session spéciale en hommage à Jean-Claude DEVILLE : L'échantillonnage équilibré. *Actes des Journées de Méthodologie Statistique*, Insee, Paris.

Tillé, Y. (2022b). On calibration and balanced sampling, Webinar in memory of Jean-Claude Deville, Statistical matching of survey data trough calibration, optimal transport and balanced sampling. *International Association of Survey Statisticians*, Webinar 17.