Editors

Roberto Benedetti

Marco Bee

Giuseppe Espa

Federica Piersimoni

Agricultural Survey Methods

WILEY

AGRICULTURAL SURVEY METHODS

Edited by

Roberto Benedetti
'G. d'Annunzio' University, Chieti-Pescara, Italy

Marco Bee
University of Trento, Italy

Giuseppe Espa
University of Trento, Italy

Federica Piersimoni
National Institute of Statistics (ISTAT), Rome, Italy

WILEY
A John Wiley and Sons, Ltd., Publication

This edition first published 2010

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

Library of Congress Cataloguing-in-Publication Data

Benedetti, Roberto, 1964-
Agricultural survey methods / Roberto Benedetti, Marco Bee, Giuseppe Espa, Federica Piersimoni.
p. cm.
Based on papers presented at the 1998, 2001, 2004 and 2007 International Conferences on Agricultural Statistics.
Includes bibliographical references and index.
ISBN 978-0-470-74371-3 (cloth)
1. Agriculture–Statistical methods. I. Bee, Marco. II. Piersimoni, Federica. III. Title.
S566.55.B46 2010
338.1021–dc22

2010000132

A catalogue record for this book is available from the British Library.

ISBN: 978-0-470-74371-3 (H/B)

Typeset in 10/12 Times-Roman by Laserwords Private Limited, Chennai, India
Printed and bound in the United Kingdom by CPI Antony Rowe, Chippenham, Wiltshire.

Agricultural Survey Methods

To Agnese and Giulio
Roberto Benedetti

To Chiara and Annalisa
Marco Bee

To Massimo, Guido and Maria Rita
Giuseppe Espa

To Renata and Oscar
Federica Piersimoni

List of Contributors

Luca Alinovi
Agricultural Development Economics Division (ESA)
FAO of the United Nations
Rome, Italy.
Luca.Alinovi@fao.org

Dale Atkinson
Research Division
National Agricultural Statistics Service (NASS)/USDA
Fairfax, VA, USA.
Dale.Atkinson@usda.gov

Bettina Baruth
IPSC-MARS, JRC
Ispra, Italy.
Bettina.baruth@jrc.ec.europa.eu

Marco Bee
Department of Economics
University of Trento
Italy.
marco.bee@unitn.it

Roberto Benedetti
Department of Business, Statistical, Technological and Environmental Sciences (DASTA)
University 'G. d'Annunzio' of Chieti-Pescara
Pescara, Italy.
benedett@unich.it

Gianni Betti
Department of Quantitative Methods
University of Siena
Italy.
betti2@unisi.it

Andrea Carfagna
Department of Statistical Sciences
University of Bologna
Italy.
andreacarfagna@virgilio.it

Elisabetta Carfagna
Department of Statistical Sciences
University of Bologna
Italy.
elisabetta.carfagna@unibo.it

Raymond L. Chambers
School of Mathematics and Applied Statistics
University of Wollongong
Australia.
ray@uow.edu.au

Denis Chartrand
Agriculture Division
Statistics Canada
Ottawa, Canada.
Denis.Chartrand@statcan.gc.ca

Arijit Chaudhuri
Applied Statistics Unit
Indian Statistical Institute
Kolkata, India.
arijitchaudhuri@rediffmail.com

Nhu Che
Australian Bureau of Agriculture and Research Economics (ABARE)
Canberra, Australia.
NChe@abare.gov.au

Jim Cotter
National Agricultural Statistics Service (NASS)/USDA, retired.

Carrie Davies
Area Frame Section of National Agricultural Statistics Service (NASS)/USDA
Fairfax, VA, USA.
Carrie_Davis@nass.usda.gov

Jacques Delincé
Institute for Prospective Technological Studies
Seville, Spain.
jacques.delince@ec.europa.eu

Marcelle Dion
Agriculture, Technology and Transportation Statistics Branch
Statistics Canada
Ottawa, Canada.
Marcelle.Dion@statcan.gc.ca

Giuseppe Espa
Department of Economics
University of Trento
Italy.
giuseppe.espa@economia.unitn.it

Pieter Everaers
Eurostat, Directorate D
External Cooperation, Communication and Key Indicators
Luxembourg.
Pieter.EVERAERS@ec.europa.eu

Danila Filipponi
Istat
National Institute of Statistics
Rome, Italy.
danila.filipponi@istat.it

Javier Gallego
IPSC-MARS, JRC
Ispra, Italy.
javier.gallego@jrc.ec.europa.eu

Giulio Ghellini
Department of Quantitative Methods
University of Siena
Italy.
ghellini@unisi.it

Antonio Giusti
Department of Statistics
University of Florence
Firenze, Italy.
giusti@ds.unifi.it

Carol C. House
Agricultural Statistics Board
National Agricultural Statistics Service
Washington, DC, USA.
Carol_House@nass.usda.gov

Philip N. Kokic
Commonwealth Scientific and Industrial Research Organisation (CSIRO)
Canberra, Australia.
phil.kokic@csiro.au

Ulf Jorner
Statistics Sweden
Stockholm, Sweden.
Ulf.Jorner@scb.se

Claude Julien
Business Survey Methods Division
Statistics Canada
Ottawa, Canada.
Claude.Julien@statcan.gc.ca

Erdgin Mane
Agricultural Development Economics Division (ESA)
FAO of the United Nations
Rome, Italy.
Erdgin.Mane@fao.org

Paul Murray
Agriculture Division
Statistics Canada
Ottawa, Canada.
Paul.Murray@statcan.gc.ca

Jack Nealon
National Agricultural Statistics Service (NASS)/USDA
Fairfax, VA, USA.
Jack.Nealon@usda.gov

Jeroen Pannekoek
Department of Methodology
Statistics Netherlands
The Hague, Netherlands.
j.pannekoek@cbs.nl (and jpnk@cbs.nl)

Federica Piersimoni
Istat
National Institute of Statistics
Rome, Italy.
piersimo@istat.it

Paolo Postiglione
Department of Business, Statistical, Technological and Environmental Sciences (DASTA)
University 'G. d'Annunzio' of Chieti-Pescara
Pescara, Italy.
postigli@unich.it

J.N.K. Rao
School of Mathematics and Statistics
Carleton University
Ottawa, Canada.
jrao@math.carleton.ca

Ray Roberts
National Agricultural Statistics Service (NASS)/USDA
Fairfax, VA, USA.
Ray.Roberts@usda.gov

Donato Romano
Department of Agricultural and Resource Economics
University of Florence
Firenze, Italy.
donato.romano@unifi.it

Vijay Verma
Department of Quantitative Methods
University of Siena
Italy.
verma@unisi.it

Frederic A. Vogel
The World Bank
Washington, DC, USA.
fvogel@worldbank.org

Ton de Waal
Department of Methodology
Statistics Netherlands
The Hague, Netherlands.
t.dewaal@cbs.nl

Anders Wallgren
Örebro University and BA Statistiksystem AB
Vintrosa, Sweden.
ba.statistik@telia.com

Britt Wallgren
Örebro University and BA Statistiksystem AB
Vintrosa, Sweden.
ba.statistik@telia.com

Introduction

The importance of comprehensive, reliable and timely information on agricultural resources is now more than ever recognized in various practical situations arising in economic, social and environmental studies. These renewable, dynamic natural resources are necessary for many countries where the growing population pressure implies the need for increased agricultural production. To improve the management of these resources it is necessary to know at least their quality, quantity and location.

In Western countries, agriculture is assuming a more and more marginal economic role in terms of its percentage contribution to GNP, but recent radical economic and social transformations have caused a renewed interest in this sector. Such interest is due not only to economic factors but also to issues related to the quality of life and to the protection of the public health. Today the food industry suffers from phenomena that originate in the primary sector, such as diseases on farms or the production of genetically modified organisms. The growing attention of consumers to the quality of food products has strongly reinforced the need to look at a single agro-food product as the result of a chain of processes linked together. In this approach, agriculture represents not only an economic sector but also the origin of the food chain, and because of this role it deserves special attention. These aspects, together with the related recovery and protection of the environment, have led to deep modifications in the data provided in this sector.

Agricultural surveys are thus conducted all over the world in order to gather a large amount of information on the classic crops, yields, livestock and other related agricultural resources. As a result, the statistics produced are so strongly conditioned by this largely diversified demand that many countries, in order to be able to comply with these requests, began to set up a complex system of surveys based on a harmonized and integrated set of information whose design, implementation and maintenance require a strong methodological effort.

Apart from the difficulties typical of business data, such as the quantitative nature of many variables and their high concentration, agricultural surveys are indeed characterized by some additional peculiarities that often make it impossible or inefficient to make use of classical solutions proposed in the literature. In particular we refer to the following:

(a) The definition of the statistical units to be surveyed is neither obvious nor unique, because the list of possible options is quite large (family, agricultural holding, household, parcel of land, point, etc.) and its choice depends not only on the

phenomenon for which we are interested in collecting the data, but also on the availability of a frame of units of sufficient quality.

(b) Typological classifications of the statistical units are very important tools to define the estimation domains and to design an efficient survey. However, harmonized hierarchical nomenclatures are usually not available for a certain definition of statistical unit, or they do exist but are so subjective that they cannot be considered as standard.

(c) The concentration of many variables is often even higher than in other business surveys.

(d) In many countries the use of the so-called 'census frames' is considered an ordinal procedure, with the obvious consequence that the list of sampling units is not updated and is very far from the target population. This has evident implications in terms of non-sampling errors due to under- or, less dangerous, over-coverage of the list.

(e) When designing a sample, the theory suggests two classical ways of using a size measure existing in the frame: a scheme with inclusion probabilities proportional to the size of the unit, and a stratification obtained through the definition of a set of threshold levels on the size variable. Nonetheless, a typical frame of agricultural units has a large amount of auxiliaries and the size of each unit is usually multivariate. In this case the solutions proposed by the methodological literature are much more complex.

(f) At least when using geographically referred units, there often exists a particular auxiliary variable requiring ad hoc procedures to be used in a sampling environment: the remotely sensed data. Remote sensing is nothing but a tool to get information about an object without being in physical contact with the object itself, and it is usually represented by digital images sensed from a satellite or an aircraft.

As far as we know, in the current literature there exists no comprehensive source of information regarding the use of modern survey methods adapted to these distinctive features of agricultural surveys.

However, the successful series of conferences on agricultural statistics, known as ICAS (International Conference on Agricultural Statistics), demonstrates that there is a broad and recognizable interest in methods and techniques for collecting and processing agricultural data. In our opinion, the remarkable number of high-quality methodological papers presented in these conferences may serve to fill this gap.

This book originates from a selection of the methodological papers presented at this set of conferences held in Washington, DC (1998), Rome (2001), Cancún (2004) and Beijing (2007). The declared aim was to develop an information network of individuals and institutions involved in the use and production of agricultural statistics.

These conferences were organized by the national statistical offices of the hosting countries – the National Agricultural Statistics Service (NASS) and United States Department of Agriculture (USDA); Italian National Statistical Institute (Istat), Agrifood and Fishery Information Service, Mexico (SIAP) and National Bureau of Statistics of China (NBS), in collaboration with international institutes such as Eurostat, FAO, OECD, UN/ECE, and ISI.

This book is an attempt to bring together the competences of academics and of experts from national statistical offices to increase the dissemination of the most recent survey methods in the agricultural sector. With this ambition in mind, the authors were asked to extend and update their research and the project was completed by some chapters on specialized topics.

Although the present book can serve as a supplementary text in graduate seminars in survey methodology, the primary audience is constituted by researchers having at least some prior training in sampling methods. Since it contains a number of review chapters on several specific themes in survey research, it will be useful to researchers actively engaged in organizing, managing and conducting agricultural surveys who are looking for an introduction to advanced techniques from both a practical and a methodological perspective.

Another aim of this book is to stimulate research in this field and, for this reason, we are aware that it cannot be considered as a comprehensive and definitive reference on the methods that can be used in agricultural surveys, since many topics were intentionally omitted. However, it reflects, to the best of our judgement, the state of the art on several crucial issues.

The volume contains 22 chapters of which the first one can be considered as an introductory chapter reviewing the current status of agricultural statistics, and the remaining 21 are divided into five parts:

I. Census, frames, registers and administrative data (Chapters 2–5). These chapters provide an overview of the basic tools used in agricultural surveys, including some practical and theoretical considerations regarding the definitions of the statistical units. Attention is then focused on the use of administrative data that in the last few years have evolved from a simple backup source to the main element in ensuring the coverage of a list of units. The opportunity to reduce census and survey costs implies growing interest, among statistical agencies, in the use of administrative registers for statistical purposes. However it requires attitudes, methods and terms that are not yet typical in the statistical tradition. The keyword is the harmonization of the registers in such a way that information from different sources and observed data should be consistent and coherent. In particular, the expensive agricultural census activities conducted periodically in every country of the world should benefit from such a radical innovation.

II. Sample design, weighting and estimation (Chapters 6–9). These chapters review advanced methods and techniques recently developed in the sampling literature as applied to agricultural units, in the attempt to address the distinctive features (c)–(e) described above. Some interesting proposals arise from the field of small-area estimation, which has received a lot of attention in recent years due to the growing demand for reliable small-area statistics needed for formulating policies and programmes. An appraisal is provided of indirect estimates, both traditional and model-based, that are used because direct area-specific estimates may not be reliable due to small-area-specific sample sizes.

III. GIS and remote sensing (Chapters 10–13). These chapters describe the use of the Geographic Information System technology as a tool to manage area- and point-based surveys. These devices are applied to carry out a wide range of operations on spatial information retrieved from many kinds of mixed sources. They

provide a detailed description of the procedures currently used in the European Union and United States to develop and sample area frames for agricultural surveys. Additionally, the usefulness of remotely sensed data as the main auxiliary variables for geographically coded units is assessed through empirical evidence, and some techniques to increase the performance of their use are proposed.

IV. Data editing and quality assurance (Chapters 14–17). These chapters deal with the classical problem of error handling, localization and correction. This is strictly connected with the issue of guaranteeing data quality, which obviously plays a central role within any statistical institute, both in strategic decisions and in daily operations. In this framework, it is evident that quality is not as much concerned with the individual data sets as with the whole set of procedures used. Some approaches to ensure data quality in collecting, compiling, analysing and disseminating agriculture data are described.

V. Data dissemination and survey data analysis (Chapters 18–22). These chapters examine some experiences in the use of statistical methods to analyse agricultural survey data. In particular, regression analysis (or some of its generalizations) is quite often applied to survey microdata to estimate, validate or forecast models formalized within agricultural economics theory. The purpose is to take into account the nature of the data analysed, as observed through complex sampling designs, and to consider how, when and if statistical methods may be formulated and used appropriately to model agricultural survey data. Web tools and techniques to assist the users to access statistical figures online are then described, for complete, safe and adequate remote statistical analyses and dissemination.

We would like to thank Daniel Berze of the International Statistical Institute for useful advice and suggestions during the starting phase of the project. Thanks are also due to Susan Barclay, Heather Kay and Richard Davies of John Wiley & Sons, Ltd for editorial assistance, and to Alistair Smith of Sunrise Setting Ltd for assistance with LaTeX.

Finally, we are grateful to the chapter authors for their diligence and support for the goal of providing an overview of such an active research field. We are confident that their competence will lead to new insights into the dynamics of agricultural surveys methods, to new techniques to increase the efficiency of the estimates, and to innovative tools to improve the timeliness and comprehensiveness of agricultural statistics.

Roberto Benedetti
Marco Bee
Giuseppe Espa
Federica Piersimoni
Pescara, Trento and Rome

November 2009

Contents