Analysis of Microdata

Rainer Winkelmann
Stefan Boes

Analysis of Microdata

With 38 Figures
and 41 Tables

Springer

Professor Dr. Rainer Winkelmann
Dipl. Vw. Stefan Boes

University of Zurich
Socioeconomic Institute
Zürichbergstrasse 14
8032 Zurich
Switzerland
E-mail: winkelmann@sts.unizh.ch
E-mail: boes@sts.unizh.ch

Cataloging-in-Publication Data

Library of Congress Control Number: 2005935030

ISBN-10 3-540-29605-0 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-29605-8 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media
springeronline.com

Printed in Germany

Cover design: Erich Kirchner
Production: Helmut Petri
Printing: Strauss Offsetdruck

SPIN 11573999 Printed on acid-free paper – 42/3153 – 5 4 3 2 1 0

Preface

The availability of microdata has increased rapidly over the last decades, and standard statistical and econometric software packages for data analysis include ever more sophisticated modeling options. The goal of this book is to familiarize readers with a wide range of commonly used models, and thereby to enable them to become critical consumers of current empirical research, and to conduct their own empirical analyses.

The focus of the book is on regression-type models in the context of large cross-section samples. In microdata applications, dependent variables often are qualitative and discrete, while in other cases, the sample is not randomly drawn from the population of interest and the dependent variable is censored or truncated. Hence, models and methods are required that go beyond the standard linear regression model and ordinary least squares. Maximum likelihood estimation of conditional probability models and marginal probability effects are introduced here as the unifying principle for modeling, estimating and interpreting microdata relationships. We consider the limitation to maximum likelihood sensible, from a pedagogical point of view if the book is to be used in a semester-long advanced undergraduate or graduate course, and from a practical point of view because maximum likelihood estimation is used in the overwhelming majority of current microdata research.

In order to introduce and explain the models and methods, we refer to a number of illustrative applications. The main examples include the determinants of individual fertility, the intergenerational transmission of secondary school choices, and the wage elasticity of female labor supply. The models presented, while chosen with economic applications in mind, should be equally relevant for other social sciences, for example, quantitative political science and sociology, and for empirical disciplines outside of the social sciences.

The book can be used as a textbook for an advanced undergraduate, a Master's or a first-year Ph.D. course on the topic of microdata analysis. In economics and related disciplines, such a course is typically offered after a first course on linear regression analysis. Alternatively, the book can also serve as a supplementary text to an applied microeconomics field course, such as those offered in the areas of labor economics, health economics, and the like. Finally, it is intended as a reference for graduate students, researchers as well as practitioners who encounter microdata in their work. The mathematical prerequisites are not very high. In particular, the use of linear algebra is minimal. On the other hand, some background in mathematical statistics is useful although not absolutely necessary.

The book includes numerous exercises. Most of the exercises do not require the use of a computer. Rather, they typically present specific empirical results, and the task is to assess the validity of the procedure in that particular context and to provide a correct interpretation of the estimated parameters. In addition, we encourage the reader to develop practical skills in applied data analysis by re-estimating the examples we discuss, using a software of choice. For this purpose, we have made the datasets employed available at our homepage www.unizh.ch/sts/, both in ASCII format and in Stata 7 format.

An earlier version of the manuscript was used in a course of the same name taught by us for several years at the economics department of the University of Zurich. We thank the participants for numerous suggestions for improvement. We are heavily indebted to Markus Lipp and Adrian Bruhin for careful proofreading, to Markus in addition for creating all the figures, and to Deborah Bowen for improving our English.

Zurich, September 2005

Rainer Winkelmann
Stefan Boes

Contents