Preface to the Second Edition

Microeconometrics Using Stata, published in December 2008, was written for Stata 10.1. *Microeconometrics Using Stata, Revised Edition*, published in January 2010, was written for Stata 11.0. This second edition is written for Stata 17.

Whereas the scope and coverage of the preceding editions were reasonably synchronized with our own *Microeconometrics: Methods and Applications* (Cambridge, 2005), this second edition has broader scope in several respects. We have attempted not only to update our previous coverage to bring it in line with newer tools in the latest edition of Stata but also to bring into the book many topics and methods that are now actively studied and increasingly used in applied microeconometrics. This coverage includes several topics, listed below, that were not covered in our 2005 text.

This second edition covers over ten years of both enhancements to Stata and developments in the methods most commonly used in empirical microeconometrics analysis. The focus of the book remains the use of linear and nonlinear regression methods for cross-sectional and short panel data. In particular, we give only short treatment to other features of Stata that are useful for data analysis such as data management, use within Stata of other programming languages such as Python, and automated document preparation. The new edition is much expanded and is split into two volumes.

The first volume, comprising chapters 1–15 and Stata and Mata appendixes, focuses on the linear regression model and provides a brief introduction to nonlinear regression models. This volume is an expanded version of chapters 1–10, 12–13, and the appendixes of the first and revised editions. In places, there is greater explanation of underlying methods, and much of the first volume is intended to be suitable for an advanced undergraduate course in addition to serving graduate students and researchers.

The second volume, comprising chapters 16–30, covers the standard nonlinear models as well as more advanced and more recent material. In addition to updated versions of chapters 14–18 of the first edition and the revised edition, the second volume includes new chapters on duration models, treatment effects in randomized control trials, treatment effects with endogenous treatments, parametric models for endogeneity and heterogeneity, spatial regression, semiparametric regression, machine learning and prediction, and Bayesian methods.

Some methods we cover are well established. Other methods we present are in areas of active research, so they may become replaced by better methods. In particular, many methods for causal analysis using observational or experimental methods are still being established and improved upon, at a remarkably rapid pace. This includes

inference for instrumental variables with weak instruments, cluster–robust inference with few clusters, treatment-effects estimation with heterogeneous treatment effects, regression discontinuity design, and causal analysis using machine learning methods. Accordingly, we plan to periodically add some supplementary material on the book's website (http://cameron.econ.ucdavis.edu/mus2).

Our target user base consists of practitioners of applied microeconometrics. This group is quite diverse in terms of familiarity with the available econometric tools. In deference to such diversity, we have chosen to separate the more advanced aspects of many topics and place them in different parts of the book. This is a challenging task because often the same material could, and in some cases should, appear in several alternative places. To assist the reader, we have provided numerous cross-references and a much lengthier subject index. The reader will benefit from checking out these connections.

Datasets and the do-files used in this book are available on the Stata Press website at https://www.stata-press.com/data/mus2.html. Any corrections to the book will be documented at https://www.stata-press.com/books/microeconometrics-stata/.

The preparation of this second edition has benefited from generous help from many sources. We thank our colleagues, coauthors, students, and many users of the previous editions for their suggested improvements, for reading parts of the book, for permission to use datasets developed in joint research, and for encouragement to proceed with the project. We have benefited from presenting some of the material in various short courses around the world and from positive feedback from readers of the earlier editions that encouraged writing this updated edition. Colin Cameron would especially like to thank Shu Shen, Takuya Ura, Oscar Jorda, Marianne Bitler, the broader econometrics and empirical microeconomics community at the University of California–Davis, and Doug Miller and Adrian Pagan. Pravin Trivedi gratefully acknowledges the support provided by the School of Economics, University of Queensland. We thank Yulia Marchenko and Nikolay Balov for very detailed comments on the Bayesian chapters, and Kristin MacDonald for a careful reading of the final draft of the book. We thank David Culwell for his excellent editing and Stephanie White for managing the LaTeX formatting and production of this book. Most especially, both authors acknowledge their debt of gratitude to David Drukker for extensive feedback on many aspects of the material in the book throughout this project, including a complete reading, as well as feedback on the substantive aspects of applying the econometric and statistical tools. Finally, we thank our respective families for their patience and understanding during the long gestation period of the evolution of this project.

Davis, CA   A. Colin Cameron
Charlottesville, VA   Pravin K. Trivedi
June 2022