

Brief contents

1 Introduction

I Continuous predictors

- 2 Continuous predictors: Linear
- 3 Continuous predictors: Polynomials
- 4 Continuous predictors: Piecewise models
- 5 Continuous by continuous interactions
- 6 Continuous by continuous by continuous interactions

II Categorical predictors

- 7 Categorical predictors
- 8 Categorical by categorical interactions
- 9 Categorical by categorical by categorical interactions

III Continuous and categorical predictors

- 10 Linear by categorical interactions
- 11 Polynomial by categorical interactions
- 12 Piecewise by categorical interactions
- 13 Continuous by continuous by categorical interactions
- 14 Continuous by categorical by categorical interactions

IV Beyond ordinary linear regression

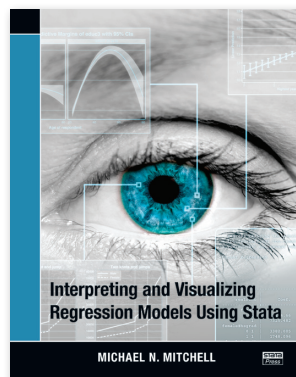
- 15 Multilevel models
- 16 Time as a continuous predictor
- 17 Time as a categorical predictor
- 18 Nonlinear models
- 19 Complex survey data

V Appendices

- A The margins command
- B The marginsplot command
- C The contrast command
- D The pwcompare command

	Coef.	Robust Std. Err.	z	P> z
female	17144.66	1389.615	12.34	0.000
female#hsgrad	0.1	3382.005	5.11	0.000
1 1	1748.096	389.3635	4.49	0.000
female#educ	86.70762	151.379	0.57	0.567
1 1	243.8081	98.9163	2.44	0.015

Interpreting and Visualizing Regression Models Using Stata



By Michael N. Mitchell

Publisher: Stata Press

Copyright: 2012

ISBN-13: 978-1-59718-107-5

Pages: 558; paperback

Price: **\$58**



For more details or to order, visit us online at
stata-press.com/books/interpreting-visualizing-regression-models.



How to contact us

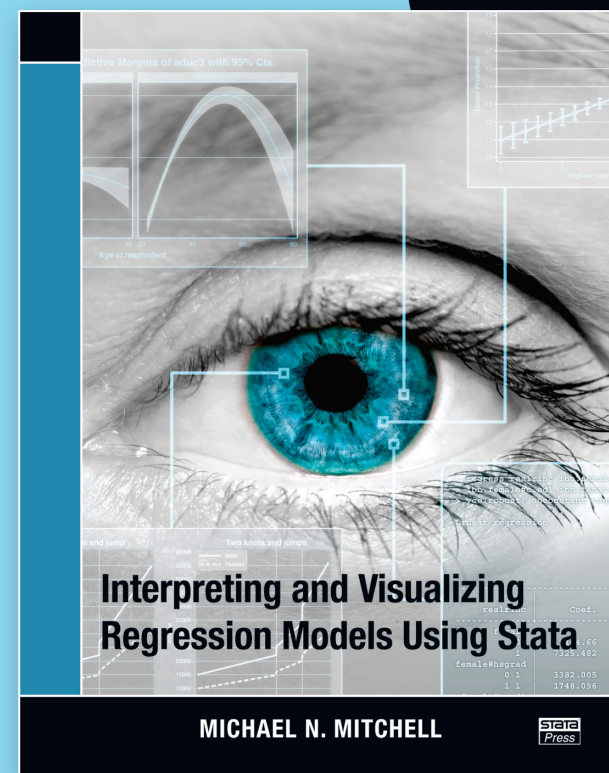
Stata Press
4905 Lakeway Drive
College Station, TX 77845-4512
USA

800-782-8272 (USA)
800-248-8272 (Canada)
979-696-4600 (Worldwide)

service@stata-press.com
stata-press.com

© Copyright 2012 StataCorp LP

Stata Press presents



About the author

Michael Mitchell is a senior statistician in disaster preparedness and response. He is also the author of *A Visual Guide to Stata Graphics* as well as *Data Management Using Stata*. Previously, he worked for 12 years as a statistical consultant and manager of the UCLA ATS Statistical Consulting Group. There, he envisioned the UCLA Statistical Consulting Resources website and wrote hundreds of webpages about Stata.

The book's audience

Interpreting and Visualizing Regression Models Using Stata applies to a wide audience of researchers, because it explores a realm of statistics not taught in the classroom: that of presenting complex models in a simple fashion. Those who would enjoy and learn from this book include:

- Statisticians who collaborate with nonstatisticians (which means most all statisticians)
- Statistical consultants
- Anyone who has ever needed to explain interactions (effect modifications) or nonlinear models
- Instructors looking to make statistical modeling more accessible
- Anyone wanting to expand his or her repertoire of modeling to include
 - » Discontinuous models, such as those based on educational milestones or policy changes
 - » Models with complex interaction terms
- Students trying to get good mental images of complex models

Comment from the Stata technical group

Michael Mitchell's *Interpreting and Visualizing Regression Models Using Stata* is a clear treatment of how to carefully present results from model-fitting in a wide variety of settings. It is a boon to anyone who has to present the tangible meaning of a complex model in a clear fashion, regardless of the audience. As an example, many experienced researchers start to squirm when asked to give a simple explanation of the practical meaning of interactions in nonlinear models such as logistic regression. The techniques presented in Mitchell's book make answering those questions easy. The overarching theme of the book is that graphs make interpreting even the most complicated models containing interaction terms, categorical variables, and other intricacies straightforward.

Using a dataset based on the General Social Survey, Mitchell starts with basic linear regression with a single independent variable, and then illustrates how to tabulate and graph predicted values. While illustrating, Mitchell focuses on Stata's `margins` and `marginsplot` commands, which play a central role in the book and which greatly simplify the calculation and presentation of results from regression models. In particular, through use of the `marginsplot` command, Mitchell shows how you can graphically visualize every model presented in the book. Gaining insight into results is much easier when you can view them in a graph rather than in a mundane table of results.

Mitchell then proceeds to more-complicated models where the effects of the independent variables are nonlinear. After discussing how to detect nonlinear effects, he presents examples using both standard polynomial terms (squares and cubes of variables) as well as fractional polynomial models, where independent variables can be raised to powers like -1 or $1/2$. In all cases, Mitchell again uses the `marginsplot` command

to illustrate the effect that changing an independent variable has on the dependent variable. Piecewise-linear models are presented as well; these are linear models in which the slope or intercept is allowed to change depending on the range of an independent variable. Mitchell also uses the `contrast` command when discussing categorical variables; as the name suggests, this command allows you to easily contrast predictions made for various levels of the categorical variable.

Interaction terms can be tricky to interpret, but Mitchell shows how graphs produced by `marginsplot` greatly clarify results. Individual chapters are devoted to two- and three-way interactions containing all continuous or all categorical variables and include many practical examples. Raw regression output including interactions of continuous and categorical variables can be nigh impossible to interpret, but again Mitchell makes this a snap through judicious use of the `margins` and `marginsplot` commands in subsequent chapters.

The first two-thirds of the book is devoted to cross-sectional data, while the final third considers longitudinal data and complex survey data. A significant difference between this book and most others on regression models is that Mitchell spends quite some time on fitting and visualizing discontinuous models—models where the outcome can change value suddenly at thresholds. Such models are natural in settings such as education and policy evaluation, where graduation or policy changes can make sudden changes in income or revenue.

This book is a worthwhile addition to the library of anyone involved in statistical consulting, teaching, or collaborative applied statistical environments. Graphs greatly aid the interpretation of regression models, and Mitchell's book shows you how.