

1 Introduction

1.1 What is this book about?

Our book shows you effective and effective ways to use regression models for categorical and count outcomes. It is a book about data analysis and is not a formal treatment of statistical models. To be effective in analyzing data, you want to spend your time thinking about substantive issues, not laboring to get your software to generate the results of interest. Accordingly, good data analysis requires good software and good technique.

Although we believe that these points apply to all data analysis, they are particularly important for the regression models that we examine. The reason is that these models are nonlinear, and consequently the simple interpretations that are possible in linear models are no longer appropriate. In nonlinear models, the effect of each variable on the outcome depends on the level of all variables in the model. Because of this nonlinearity, which we discuss in detail in chapter 3, no method of interpretation can fully describe the relationships among the independent variables and the outcome. Rather, a series of *postestimation* explorations are needed to uncover the most important aspects of these relationships. In general, if you limit your interpretations to the standard output of estimated slope coefficients, your interpretation will usually be incomplete and sometimes it will even be misleading.

In the linear regression model (LRM), most of the work of interpretation is complete once the estimates are obtained. You simply read off the coefficients, which can be interpreted as: “For a unit increase in x_k , y is expected to increase by β_k units, holding all other variables constant.” In nonlinear models, such as logit or negative binomial regression, additional computations are necessary after the estimates are obtained. Indeed, when interpreting nonlinear models, most of the work involves sometimes complex postestimation analyses, which are the focus of our book.

To make these computations, we use Stata’s postestimation commands along with commands that we have written. Without these commands, the computations are time consuming, error prone, and tedious. All in all, it is not fun work and it is tempting to limit your analyses to an uninformative table of parameter estimates. Fortunately, the commands we discuss in this book make sophisticated postestimation analysis routine and even enjoyable. While these analyses can take a lot of work, our commands reduce the tedium so that the data analyst can focus on substantive issues.

1.2 Which models are considered?

Regression models analyze the relationship between an explanatory variable and an outcome variable while controlling for the effects of other variables. The linear regression model (LRM) is probably the most commonly used regression model in the social sciences. A key advantage of the LRM is the ease of interpreting results. Unfortunately, this model applies only to cases in which the dependent variable is continuous. Using the LRM when it is not appropriate may produce coefficients that are biased and inconsistent, and there is nothing advantageous about the simple interpretation of results that are incorrect. Fortunately, many appropriate models exist for categorical outcomes, and these models are the focus of our book. We cover models for four kinds of dependent variables.

Binary outcomes have two values, such as whether a citizen voted in the last election, whether a patient was cured after receiving some medical treatment, or whether a respondent attended college. The regression models and commands we consider include binary logit (`logit`) and binary probit (`probit`).

Ordinal outcomes have more than two categories that are assumed to be ordered on a single, underlying dimension. For example, a survey might ask if you would be “very likely”, “somewhat likely”, or “not at all likely” to take a new subway to work, or if you agree with the president on “all issues”, “most issues”, “some issues”, or “almost no issues”. We focus on the ordered logit (`ologit`) and ordered probit (`oprobit`) models, but also consider the sequential logit model (`seqlogit`), stereotype logistic regression (`slogit`), and the generalized ordered logit (`gologit2`) which is also appropriate for nominal outcomes.

Nominal outcomes also have more than two categories but the categories are not ordered. Examples include the mode of transportation a person takes to work (e.g., bus, car, train) or an individual’s employment status (e.g., employed, unemployed, out of the labor force). The primary model we consider is the multinomial logit model (`mlogit`), along with its counterpart the multinomial probit model with uncorrelated errors (`mprobit`). We also review the related conditional logit model (`clogit` and `asclogit`), the alternative-specific multinomial probit with correlated errors (`asmprobit`), and the rank-ordered logit model (`rologit`).

Finally, count variables count the number of times something has happened, such as the number of articles written by a scientist or the number of patents a biotechnology company has obtained. We begin with the Poisson regression model (`poisson`), followed by the negative binomial regression model (`nbreg`), zero-truncated Poisson and negative binomial model (`tpoisson` and `tnbreg`), the hurdle regression model, and end with an examination of the zero-inflated Poisson and negative binomial models (`zip` and `zinb`).

Although this book covers many models for different types of outcomes, they are all models for cross-sectional data. We do not consider models for survival or event-history data, even though Stata has a powerful set of commands for dealing with these data. We recommend [Cleves et al. \(2010\)](#) and the *Survival Analysis Reference Manual*

for more information on these models. Likewise, we do not consider any models for panel or other multilevel data, even though Stata contains commands for fitting these models. For additional information, see [Rabe-Hesketh and Skrondal \(2012\)](#) and the *Longitudinal-Data/Panel-Data Reference Manual*.

1.3 Whom is this book for?

We expect that readers of this book will vary considerably in their knowledge of both statistics and Stata. With this in mind, we have tried to structure the book to accommodate the diversity of our audience. Minimally, however, we assume that readers have a solid familiarity with the linear regression model and that they are comfortable using the basic features of the operating system of their computer. Although we have provided sufficient information about each model so that you can read each chapter without prior exposure to the models discussed, we strongly recommend that you do not use this book as your sole source of information on the models (section 1.8 recommends more readings). Our book will be most useful if you have already studied the models considered or are studying these models in conjunction with our book.

Ideally you are running Stata 13 or later. Most of our examples will, however, run in Stata 11 and 12. If you are using a version of Stata earlier than Stata 11, we suggest that you use the second edition of our book ([Long and Freese 2006](#)). However, with the powerful new features in Stata 13 and the new methods of interpretation in the current edition, we hope you decide instead to upgrade your software. To make the most out of the book, you will need access to the Internet to download our commands, datasets, and sample programs (see section 1.5 for details on obtaining these). For information about obtaining Stata, see the StataCorp website at <http://www.stata.com>.

1.4 How is the book organized?

Chapters 2 through 4 introduce materials that are essential for working with the models we present in the later chapters:

Chapter 2: Introduction to Stata reviews the basic features of Stata that are necessary to get new or inexperienced users up and running with the program. New users should work through the brief tutorial that we provide in section 2.18. This introduction is by no means comprehensive, so we include information on how to learn more about using Stata. Those who are familiar with Stata can skip this chapter, although even these readers might benefit from quickly reading it.

Chapter 3: Estimation, testing, and fit reviews Stata commands for fitting models, testing hypotheses, and computing measures of model fit. While those who regularly use Stata for regression modeling might be familiar with much of this material, we suggest at least a quick review of the material. Most importantly, you should read our detailed discussion of factor-variable notation, which was in-

roduced in Stata 11. Understanding how to use factor variables is essential for the methods of interpretation presented in the book.

Chapter 4: Methods of interpretation is an overview of various approaches to interpreting regression models. We introduce the `margins` command that is part of official Stata and the `mtable`, `mgen`, and `mchange` that are part of `SPost13`. This chapter is essential background before proceeding to Part II. It is important to study this chapter carefully even if you are an advanced user. Readers who are new to Stata are likely to find that this chapter has more detail than you need initially. As we proceed, we suggest which sections one may wish to only skim on first reading.

Part II covers regression models for a different types of outcomes.

Chapters 5 and 6: Models for binary outcomes begins with an overview of how the binary logit and probit models are derived and how they can be fitted. After the model has been fitted, we show how to test hypotheses, compute residuals and influence statistics, and calculate scalar measures of model fit. Chapter 6 uses postestimation commands that assist in interpretation using predicted probabilities, discrete and marginal change in the predicted probabilities, and, for the logit model, odds ratios. Because binary models provide a foundation on which many models for other kinds of outcomes are derived, and because these two chapters provide more detailed explanations of common tasks than later chapters do, we recommend reading this chapter carefully even if you are interested mainly in another type of outcome.

Chapter 7: Models for ordinal outcomes presents the ordered logit and ordered probit models. We show how these models are fitted and how to test hypotheses about coefficients. We also consider tests of a key assumption of both models, known as the parallel regression assumption. For interpreting results, we discuss similar methods as those described in chapter 6, as well as interpretation in terms of a latent dependent variable. Methods of interpretation using predicted probabilities apply directly to models for nominal outcomes, so it is useful to familiarize yourself with these methods before proceeding to chapters 8. This chapter also discusses the implications of assuming that an ordinal model is appropriate for your outcome and recommend that you use models for nominal outcomes as part of your evaluation of ordinal models.

Chapter 8: Models for nominal outcomes focuses on the multinomial logit model. We show how to test hypotheses that involve multiple coefficients and discuss tests a key assumption known as the independence of irrelevant alternatives assumption. Methods of interpretation using predictions are *identical* to those for ordinal models. Interpretation using odds ratio is a simple extension of methods introduced in chapter 6, although the multinomial logit model's many parameters makes the process of interpretation much more complicated. To deal with this complexity, we present a graphical method for summarizing the parameters. The multinomial

probit model without correlated errors is discussed briefly, and then the multinomial logit model is used to explain the stereotype logit model. This model, which is often used with ordinal outcomes, also has applications with nominal outcomes. These models assume case-specific independent variables, which means each independent variable has one value for each observation. We end the chapter with a short review of models that include also alternative-specific data, in which some variables vary over the alternatives for each individual, such as an individual's similarity to each candidate in an election. We consider the conditional logit model and the alternative-specific multinomial probit model, the latter of which allows correlations between alternative-specific error terms. Last, we present the rank-ordered logistic regression model, which can be used when you have information about the ranking of outcomes as opposed to only information about the selected or most preferred outcome.

Chapter 9: Models for count outcomes begins with the Poisson and negative binomial regression models, including a test to determine which model is appropriate for your data. We also show how to incorporate differences in exposure time into parameter estimation. Next we consider interpretation for changes in the predicted rate and changes in the predicted probability of observing a given count. The rest of the chapter deals with models that address problems associated with having too many zeros relative to what the model predicts or no zeros at all. We start with zero-truncated models for which zeros are missing from the outcome variable, perhaps due to the way the data were collected. We then merge a binary model and a zero-truncated model to create the hurdle model. The rest of the chapter considers fitting and interpreting zero-inflated count models, which are designed to account for the many zero counts often found in count outcomes.

1.5 The SPost software

From our point of view, one of the best things about Stata is how easy it is to add your own commands. This means that if Stata does not have a command you need or some command does not work the way you like, you can program the command yourself, and it will work as if it were part of official Stata. We have created a suite of programs, referred to collectively as SPost13 (for Stata Postestimation version 13 commands), for the postestimation interpretation of regression models. These commands must be installed before you can try the examples in later chapters.

If you have used SPost before – read this! For this book, we completely rewrote our earlier SPost commands, which we will refer to as SPost9. If you have the `spost9_ado` package installed on your computer, you should uninstall it (details below) before you install the `spost13_ado` package.

To get the most out of this book, you need to try each method using the both official

Stata commands and our `SPost13` commands. Ideally you are running Stata 13 or later, and some methods we present require this. If you are running Stata 11 or Stata 12, most of our examples will work, but a few valuable features will not be available. Before we discuss how to install our commands and update your software, we have suggestions for new Stata users, those with earlier versions of Stata, and those who used `SPost9`:

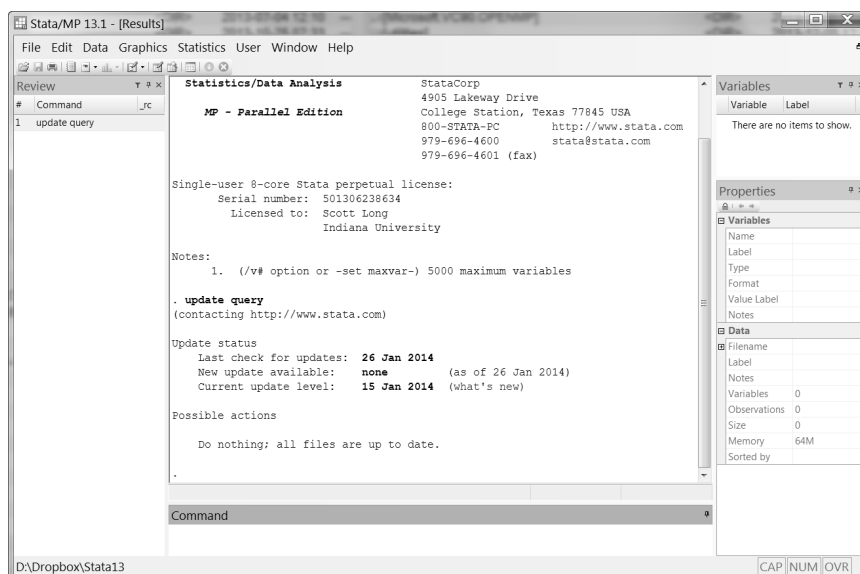
Advice to new Stata users If you have never used Stata, you might find the instructions in this section to be confusing. It might be easier if you skim the material now and return to it after you have read the the introduction to Stata in chapter 2.

If you are using Stata 10 and earlier The `SPost13` commands used in this book will not with run with Stata 10 and before. You can use `SPost9` contained in the `spost9_ado` package that is described in the second edition of our book (Long and Freese 2006). However, if you are investing the time to learn these methods, we think you are much better off upgrading your software so that you can use `spost13_ado`.

If you are using an earlier version of SPost Before using the `spost13_ado` package, you must uninstall any earlier versions of `SPost` such as the `spostado` package or the `spost9_ado` package. `SPost13` replaces our earlier `prvalue`, `prtab`, `prgen`, `prchange`, and `praccum` commands with the more powerful `mtable`, `mgen`, `mchange`, and `mlistat` commands that are based on Stata's remarkable `margins` command, which did not exist when the previous edition of this book was written. If you want to use our new commands, but also want access the `SPost9` commands, you can install the `spost9_legacy` package. Details are given below.

1.5.1 Updating Stata

Before you install `SPost13`, we strongly recommend that you update your version of Stata and the official Stata ado-files. This doesn't mean update from Stata 11 to Stata 13, but rather making sure you have the latest updates to whatever version of Stata you are running. You should do this even if you have just installed Stata because the CD or download that you received might not have the latest changes to the program. While online and in Stata, you can update Stata by selecting **Check for Updates** from the **Help** menu or equivalently you can enter the command `update query`. When we did this, Stata showed:



This screen tells you the current update status and whether you should update. Our installation is up to date. If it wasn't, Stata would provide instructions on how to update our files.

1.5.2 Installing SPost13

We begin with a little bit of background of what happens when you install user-written commands in Stata. The good news is that once they are installed, these commands behave just like commands in official Stata. Programs that add commands to Stata are contained in files that end in the extension `.ado` which stands for “automatic do-files”. For example, `listcoef.ado` is the file that contains the command `listcoef`. When you type `listcoef` at the command line, Stata automatically runs `listcoef.ado`. When this happens, it will seem just like other commands in official Stata. The `ado` files, along with supplementary files that might have other suffixes, are included as part of a *package*. In Stata, a package is essentially a list of the files along with instructions on how to install them on your computer. To install SPost13, you install the `spost13.ado` package. We consider two methods for installation, but first we explain how to uninstall the `spost9.ado` package.

Uninstalling SPost9

To determine if you have installed this package, you can run the command `ado` and see if `spost9.ado` is listed. If it is, you can uninstall it by running the command:

```
. ado uninstall spost9_ado
```

```

package spost9_ado from http://www.indiana.edu/~jslsoc/stata
Distribution-date: 25Jul2012

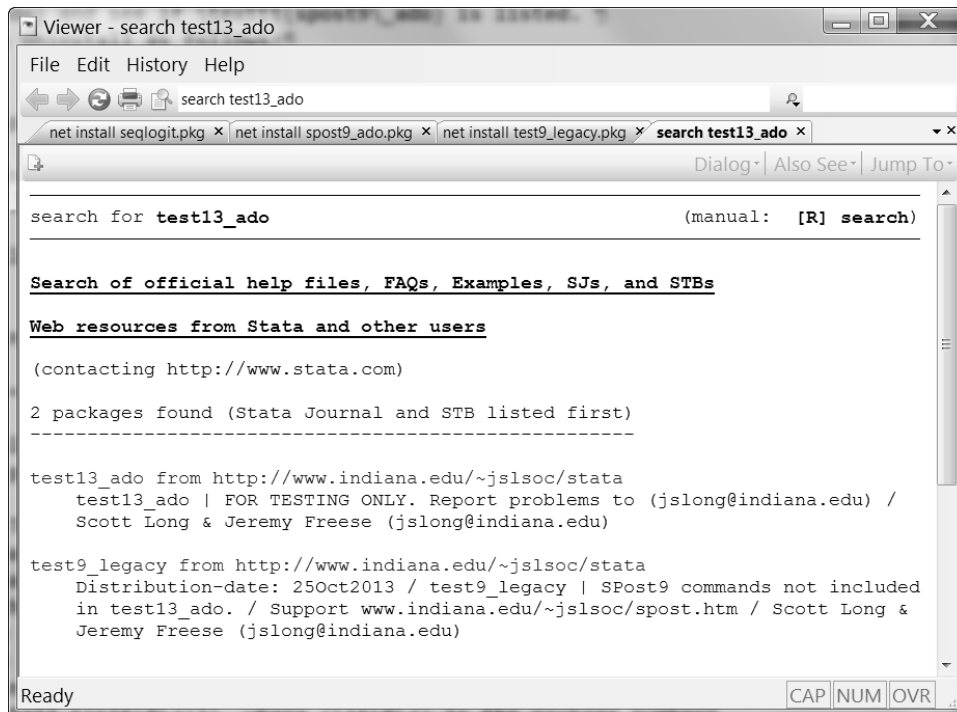
(package uninstalled)

```

Installing SPost13 using search

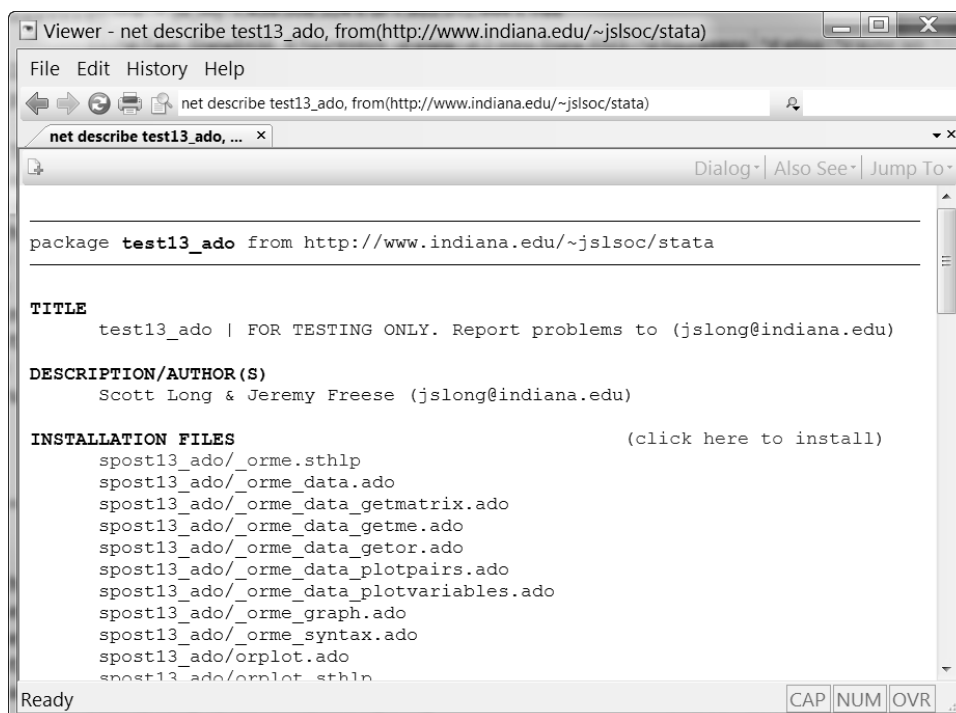
The `search word` command searches an online database that StataCorp uses to keep track of user-written additions to Stata. Typing `search spost13_ado` opens a Viewer window such as following:

Lisa: Please get a new screen capture when search `spost13_ado` is run.



If you click on `spost13_ado` from <http://www.indiana.edu/~jslsoc/stata>, which should be shown in blue, a new Viewer window opens that looks something like this:

Lisa: Please get a new screen capture.



Select **(click here to install)** and Stata will attempt to install the package. After a delay during which files are downloaded, Stata will respond with one of the following messages:

- installation complete** means that the package has been successfully installed and that you can now use the commands. Just above the “installation complete” message, Stata tells you the directory where the files were installed.
- all files already exist and are up to date** means that your system already has the latest version of the package. You do not need to install the package since it was already installed.
- the following files exist and are different** indicates that your system already has files with the same names as those in the package being installed and that the existing files differ from those in the package. The names of those files are listed and you are given several options. Assuming that the files listed are earlier versions of our programs, you should select the option “Force installation replacing already-installed files”. This might sound ominous, but it is not. Because the files on our website are the latest versions, you want to replace your current files with these new files. After you accept this option, Stata updates your files to newer versions.

cannot write in directory *directory-name* means that you do not have write privileges to the directory where Stata wants to install the files. Usually, this occurs only when you are using Stata on a network. Then we recommend that you contact your network administrator and ask if our commands can be installed using the instructions given above. If you cannot wait for a network administrator to install the commands or to give you the needed write access, you can install the programs to any directory where you have write permission, including a flash drive or your personal directory on a network. For example, suppose that you want to install SPost13 to your directory called `d:\username` (which can be any directory where you have write access). You should use the following commands:

```
. cd d:\username
d:\username
. mkdir ado
. sysdir set PERSONAL "d:\username\ado"
. net set ado PERSONAL
. net search spost
(contacting http://www.stata.com)
```

Then follow the installation instructions that we provided earlier for installing SPost13. If you get the error “could not create directory” after typing `mkdir ado`, you probably do not have write privileges to the directory.

If you install ado-files to your own directory, each time you begin a new session you must tell Stata where these files are located. You do this by typing `sysdir set PERSONAL directory`, where *directory* is the location of the ado-files you have installed. For example,

```
. sysdir set PERSONAL d:\username\ado
```

Installing SPost13 using net install

You can also install the the `spost13_ado` package entirely from the Command window. If the method listed above doesn't work, the following steps might. While online, enter

```
. net from http://www.indiana.edu/~jslsoc/stata/
```

The available packages will be listed in the Results window. You can click on `spost13_ado` and follow instructions, or you can type

```
. net install spost13_ado
```

`net get` can be used to download supplementary files (e.g., datasets, sample do-files) from our website. For example, to download the package `spost13_do` (discussed below), type

```
. net get spost13_do
```

These files are downloaded to the current working directory (see chapter 2 for a full discussion of the working directory).

1.5.3 Uninstalling SPost13

Stata keeps track of the packages that it has installed, which makes them easy to uninstall in the future. If you want to uninstall our commands, simply type `ado uninstall spost13_ado`. When things don't seem to work right, one of the first things we suggest is to uninstall `spost13_ado` and then re-install it.

1.6 Sample do-files and datasets

While we hope you will soon be using the methods in the book with your own data, we think it is valuable to reproduce our examples and modify them to try new things. To facilitate this, we have written the `spex` command to make it easy to load our data and run our baseline models and created the `spost13.do` package that lets you download the data and do-files for each chapter.

1.6.1 Installing the `spost13.do` package

The `spost13.do` package contains the datasets used in the book along with do-file that reproduce most of the analyses. The do files have names like `rm3ch9-count.do` and contain comments that link the commands to sections of the book. To download these files to your working directory, while in Stata and connected to the internet, type `search spost13.do`. This opens a Viewer window. If you click on the blue link `spost13.do` from <http://www.indiana.edu/~jslsoc/stata> and a new Viewer window opens. Follow the instructions to download the files.

In the examples shown throughout the book, we assume the commands are being run in a working directory in which the `spost13.do` package has been installed.

1.6.2 Using `spex` to load data and run examples

Experimenting with the postestimation commands that we discuss requires that you have fitted the appropriate model. In our examples, we show you how to open a dataset and fit models as you would when you were working with your own data. Accordingly, we begin with a `use` command to load the data and then use an estimation command, such as `logit`, to fit the model. To make it simpler for you to experiment with the methods in later chapters, we have written the command `spex` (Stata postestimation examples). If you type `spex logit`, for example, it will automatically load the data and fit the model that serves as our main logit example. Typing `spex commandname` will produce our primary example for that estimation command. Or, you can specify the name of any dataset that we use, `spex datasetname`, and `spex` will load those data but not fit any model. By default, `spex` looks for the dataset on our web site. If it does not find it there, it will look in the current working directory and all the directories where Stata searches for ado-files. For more information, type `help spex`.

The running examples in this edition of the book are different than those used in the second edition. If you want to run the earlier examples, use the `spex9` command. For example, `spex9 logit` runs the logit example that was used with `spost9_ado`.

1.7 Getting help with SPost

Things don't always work as intended and commands that we say will work might not. This section talks about what to do when that happens. First, there are some things that we ask you to do. If they don't resolve the problem, you can contact us. Before you do, please read section this section carefully. Thanks.

1.7.1 What if an SPost command does not work?

This section assumes that you have installed SPost13 but that some or all of the commands do not work. Here are some things to consider:

1. Make sure that Stata is properly installed and up to date. Typing `verinst` will verify that Stata has been properly installed. Typing `update query` will tell you if the version you are running is up to date and what you need to type to update it. If you are running Stata over a network, your network administrator may need to do this for you. See [U] **28 Using the Internet to keep up to date** and [R] **update**.
2. Make sure SPost13 is up to date. Type `adoupdate`, `update` or `uninstall` then `reinstall`. The problem you encountered might have been fixed since you installed `spost13_ado`.
3. If you get the error message `unrecognized command`, there are several possibilities.
 - a. If the command used to work, you could be working on a different computer or on a different station in a computer lab. Because user-written ado-files work seamlessly in Stata, you might not realize that these programs must be installed on each machine that you use.
 - b. If you sent a do-file using SPost13 commands to another person who cannot get the commands to work, let them know that they need to install SPost13. If they are using SPost9, your do-file won't work.
 - c. If you get the error message `unrecognized command: strangename` after typing one of our commands, where *strangename* is not the name of the command that you typed, it means that Stata cannot find an ancillary ado-file that the command needs. We recommend that you `uninstall` the `spost13_ado` package and then `reinstall` it.
4. If you get an error message that you do not understand, click on the blue return code beneath the error message for more information about the error.

5. Often what appears to be a problem with one of our commands is actually a mistake the user has made. We know this because we make these mistakes, too. For example, make sure that you are not using `=` when you should be using `==`.
6. Because our commands work after you have fitted a model, make sure that there were no problems with the last model fitted. If Stata was not successful in fitting your model, our commands will not have the information needed to operate properly. Our commands should trap such errors, but sometimes don't.
7. Irregular value labels can cause commands to fail. To the extent possible, we recommend using labels that have fewer than eight characters and contain no spaces or special characters other than underscores (`_`). If you are having problems and your variables do not meet this standard (especially the labels for your dependent variable), try changing your value labels with the `label` command (details are given in section 2.15).
8. Unusual values of the outcome categories can also cause problems. For ordinal or nominal outcomes, some of our commands require that all the outcome values be integers between 0 and 99. The behavior of some official Stata commands can also be confusing when unusual values are used. For these types of outcomes, we strongly recommend using consecutive integers starting with 1.

1.7.2 Getting help from the authors

If you have tried everything we recommended in section 1.7.1 and you are still encountering errors, the next step is to contact us. We hear from hundreds of readers and do our best to help. To make this easier for us and to increase the chances that we will be able to help, please follow the suggestions in this section carefully. We encourage you to start by taking a few minutes to review William Gould's blog entry "How to successfully ask a question on Statalist" (<http://blog.stata.com/2010/12/14/how-to-successfully-ask-a-question-on-statalist/>). His advice will increase your chances of getting your questions answered, either from the Statalist, from us, or elsewhere. And, in our experience, in the process of carefully preparing your question you will find the solution yourself.

Here are things that you can do to make it easier for us to answer your questions, which will also increase the odds that you will get a prompt answer.

1. Check <http://www.indiana.edu/~jlsoc/spost.htm> and http://www.indiana.edu/~jlsoc/spost_help.htm for advice on what to try before contacting us. There might be recent information that solves your problem.
2. Make sure that Stata and `spost13_ado` are up-to-date. We keep repeating this because it is the most common solution to problems we hear about.
3. Look at the sample files in the `spost9_do` package. It is sometimes easiest to figure out how to use a command by seeing how others use it. Try what you are doing on one of our datasets and see if you can reproduce it.
4. If you still have a problem, send us the information described in the next section.

What we need in order to help you

To solve your problem we need to be able to reproduce it. Simply describing the problem rarely is sufficient. Please send us:

1. A do-file that reproduces the problem.
 - a. Include your name, e-mail address, and a description of the problem.
 - b. Begin with the commands `about` and `spost13which`, which tell us which versions of software you are using.
 - c. Include the results of `summarize` used in the analysis (not all variables in the dataset) and results of `tabulate` for categorical variables.
 - d. Include only the commands needed to reproduce or explain the problems. Remove all unrelated commands.
 - e. Remove all references to specific directories such as `use c:\data\project3\sample.dta` or `log using c:\data\project3\problem, text`. Or computers won't have your directory structure so your do-file won't run on our computers. The do-file should read the data from the working directory and save the log file to the working directory.
2. The dataset used by the do-file. This should be a small dataset extracted from the full dataset you are using. Only send the variables used in your do-file and ideal a dataset with only a subset of your observations (assuming, of course, that the error is reproduced with the smaller sample).
3. A plain text log file showing the error. Do not send the log in SMCL format. To avoid this, add the `text` option to your `log` command, such as `log using myproblem.log, text replace`.

Send this information to jslong@indiana.edu or jfreese@northwestern.edu. Here is an example of what a do-file might look like:

```
capture log close
log using yourname.log, text replace
* mtable generates a variable not found error.
* scott long - jslong@indiana.edu - 2014-05-09
about
whichspost13
use jslong-error.dta
logit y x1 x2
* the following command causes the error
mtable x1, at(x2=(1(1)3))
log close
```

1.8 Where can I learn more about the models?

There are many valuable sources for learning more about the regression models that are covered in this book. We recommend the following:

- Cameron, A. C. and P. K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press. This is an excellent introduction to the methods and models discussed in this book, as well as models for panel data.
- Cameron, A. C. and P. K. Trivedi. 2010. *Microeconometrics Using Stata, Revised edition*. College Station, TX: Stata Press. This companion to (Cameron and Trivedi 2005) shows how to use Stata for cross-section and panel models.
- Cameron, A. C. and P. K. Trivedi. 2013. *Regression Analysis of Count Data, 2nd edition*. Cambridge: Cambridge University Press. This is a definitive reference on count models.
- Hardin, J. and J. Hilbe. 2012. *Generalized Linear Models and Extensions, 3rd edition*. College Station, TX: Stata Press. This is a thorough review of the generalized linear model or GLM approach to modeling and includes detailed information on using these models with Stata.
- Hosmer, D. W., and S. Lemeshow, and R. X. Sturdivant. 2013. *Applied Logistic Regression, 3rd edition*. New York: Wiley. This book, written primarily for biostatisticians and medical researchers, considers logit models for binary, ordinal, and nominal outcomes. Often the authors discuss how their recommendations can be executed using Stata.
- Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage. This book provides more details about the models discussed in our book.
- Train, K. 2009. *Discrete Choice Methods with Simulation, 2nd edition*. Cambridge: Cambridge University Press. This is a thorough review of models for a wide range of models for discrete choice and includes details on new methods of estimation using simulation.
- Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data, 2nd edition*. Cambridge, MA: MIT Press. This is a comprehensive review of econometric methods for cross-sectional and panel data.