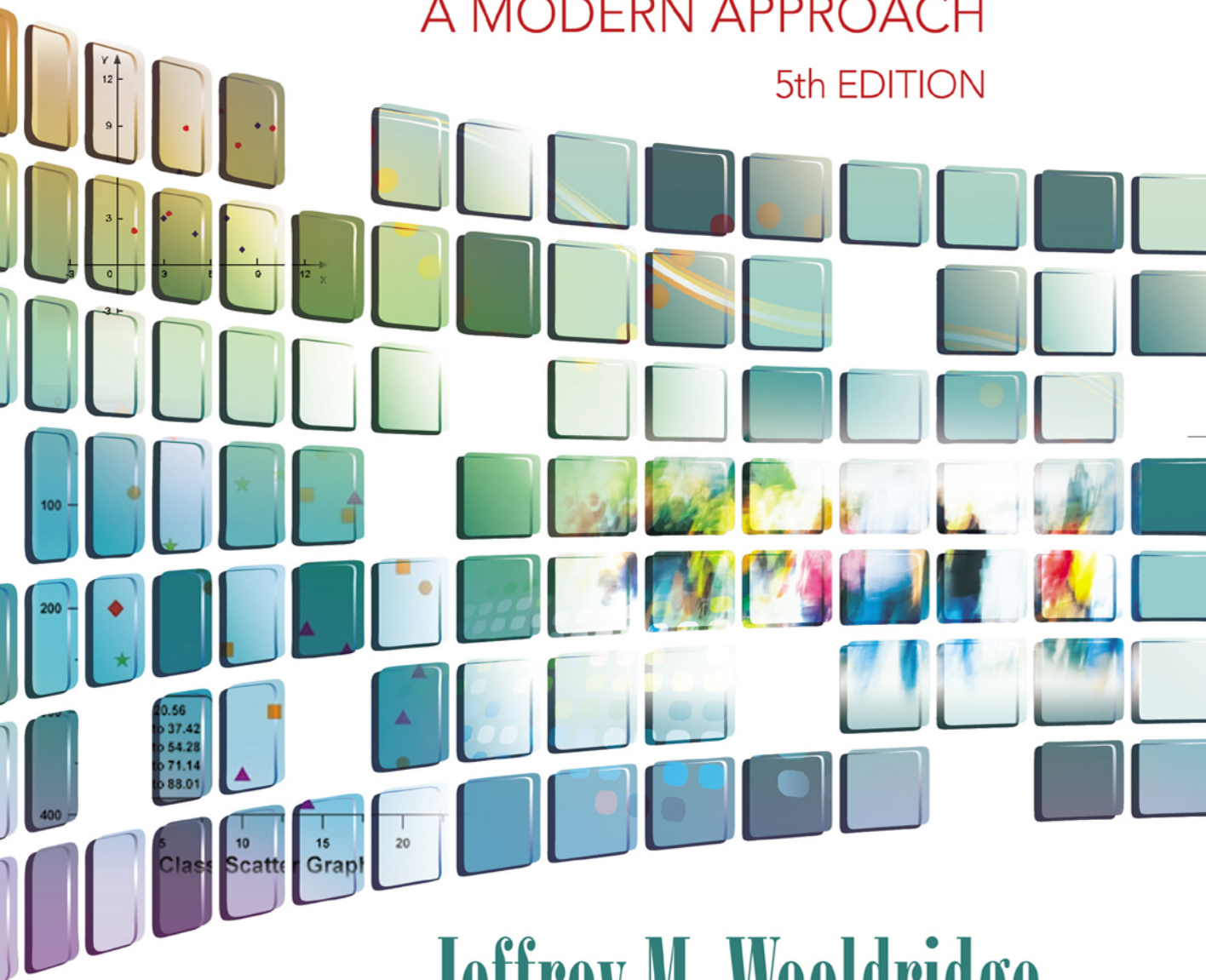


INTRODUCTORY Econometrics

A MODERN APPROACH

5th EDITION



Jeffrey M. Wooldridge

Introductory Econometrics: A Modern Approach, Fifth Edition
Jeffrey M. Wooldridge

Senior Vice President, LRS/Acquisitions & Solutions Planning: Jack W. Calhoun

Editorial Director, Business & Economics: Erin Joyner

Editor-in-Chief: Joe Sabatino

Executive Editor: Michael Worls

Associate Developmental Editor: Julie Warwick

Editorial Assistant: Libby Beiting-Lipps

Brand Management Director: Jason Sakos

Market Development Director: Lisa Lysne

Senior Brand Manager: Robin LeFevre

Senior Market Development Manager: John Carey

Content Production Manager: Jean Buttrom

Rights Acquisition Director: Audrey Pettengill

Rights Acquisition Specialist, Text/Image: John Hill

Media Editor: Anita Verma

Senior Manufacturing Planner: Kevin Kluck

Senior Art Director: Michelle Kunkler

Production Management and Composition: PreMediaGlobal

Internal Designer: PreMediaGlobal

Cover Designer: Rokusek Design

Cover Image: © Elena R/Shutterstock.com; Milosz Aniol/Shutterstock.com

© 2013, 2009 South-Western, Cengage Learning

ALL RIGHTS RESERVED. No part of this work covered by the copyright herein may be reproduced, transmitted, stored, or used in any form or by any means graphic, electronic, or mechanical, including but not limited to photocopying, recording, scanning, digitizing, taping, web distribution, information networks, or information storage and retrieval systems, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the publisher.

For product information and technology assistance, contact us at
Cengage Learning Customer & Sales Support, 1-800-354-9706

For permission to use material from this text or product,
submit all requests online at **www.cengage.com/permissions**
Further permissions questions can be emailed to
permissionrequest@cengage.com

Library of Congress Control Number: 2012945120

ISBN-13: 978-1-111-53104-1

ISBN-10: 1-111-53104-8

South-Western

5191 Natorp Boulevard
Mason, OH 45040
USA

Cengage Learning products are represented in Canada by
Nelson Education, Ltd.

For your course and learning solutions, visit **www.cengage.com**

Purchase any of our products at your local college store or at our preferred
online store **www.cengagebrain.com**

Printed in the United States of America
1 2 3 4 5 6 7 16 15 14 13 12

TABLE 19.3 OLS Results. Dependent Variable: Participation Rate			
Independent Variables	(1)	(2)	(3)
<i>mrte</i>	.156 (.012)	.239 (.042)	.218 (.342)
<i>mrte</i> ²	—	−.087 (.043)	−.096 (.073)
<i>log(emp)</i>	−.112 (.014)	−.112 (.014)	−.098 (.111)
<i>log(emp)</i> ²	.0057 (.0009)	.0057 (.0009)	.0052 (.0007)
<i>age</i>	.0060 (.0010)	.0059 (.0010)	.0050 (.0021)
<i>age</i> ²	−.00007 (.00002)	−.00007 (.00002)	−.00006 (.00002)
<i>sole</i>	−.0001 (.0058)	.0008 (.0058)	.0006 (.0061)
<i>constant</i>	1.213 (.051)	.198 (.052)	.085 (.041)
<i>industry dummies?</i>	no	no	yes
Observations	3,784	3,784	3,784
<i>R</i> -squared	.143	.152	.162

© Cengage Learning, 2013

Note: The quantities in parentheses below the estimates are the standard errors.

Summary

In this chapter, we have discussed the ingredients of a successful empirical study and have provided hints that can improve the quality of an analysis. Ultimately, the success of any study depends crucially on the care and effort put into it.

Key Terms

Data Mining

Internet

Misspecification Analysis

Online Databases

Online Search Services

Sensitivity Analysis

Spreadsheet

Text Editor

Text (ASCII) File

Sample Empirical Projects

Throughout the text, we have seen examples of econometric analysis that either came from or were motivated by published works. We hope these have given you a good idea about the scope of empirical analysis. We include the following list as additional examples of questions

that others have found or are likely to find interesting. These are intended to stimulate your imagination; no attempt is made to fill in all the details of specific models, data requirements, or alternative estimation methods. It should be possible to complete these projects in one term.

- 1 Do your own campus survey to answer a question of interest at your university. For example: What is the effect of working on college GPA? You can ask students about high school GPA, college GPA, ACT or SAT scores, hours worked per week, participation in athletics, major, gender, race, and so on. Then, use these variables to create a model that explains GPA. How much of an effect, if any, does another hour worked per week have on GPA? One issue of concern is that hours worked might be endogenous: it might be correlated with unobserved factors that affect college GPA, or lower GPAs might cause students to work more.

A better approach would be to collect cumulative GPA prior to the semester and then to obtain GPA for the most recent semester, along with amount worked during that semester, and the other variables. Now, cumulative GPA could be used as a control (explanatory variable) in the equation.

- 2 There are many variants on the preceding topic. You can study the effects of drug or alcohol usage, or of living in a fraternity, on grade point average. You would want to control for many family background variables, as well as previous performance variables.
- 3 Do gun control laws at the city level reduce violent crimes? Such questions can be difficult to answer with a single cross section because city and state laws are often endogenous. [See Kleck and Patterson (1993) for an example. They used cross-sectional data and instrumental variables methods, but their IVs are questionable.] Panel data can be very useful for inferring causality in these contexts. At a minimum, you could control for a previous year's violent crime rate.
- 4 Low and McPheters (1983) used city cross-sectional data on wage rates and estimates of risk of death for police officers, along with other controls. The idea is to determine whether police officers are compensated for working in cities with a higher risk of on-the-job injury or death.
- 5 Do parental consent laws increase the teenage birthrate? You can use state level data for this: either a time series for a given state or, even better, a panel data set of states. Do the same laws reduce abortion rates among teenagers? The *Statistical Abstract of the United States* contains all kinds of state-level data. Levine, Trainor, and Zimmerman (1996) studied the effects of abortion funding restrictions on similar outcomes. Other factors, such as access to abortions, may affect teen birth and abortion rates.

There is also recent interest in the effects of "abstinence-only" sex education curricula. One can again use state-level panel data, or maybe even panel data at the school district level, to determine the effects of abstinence-only approaches to sex education on various outcomes, including rates of sexually transmitted diseases and teen birth rates.

- 6 Do changes in traffic laws affect traffic fatalities? McCarthy (1994) contains an analysis of monthly time series data for the state of California. A set of dummy variables can be used to indicate the months in which certain laws were in effect. The file TRAFFIC2.RAW contains the data used by McCarthy. An alternative is to obtain a panel data set on states in the United States, where you can exploit variation in laws across states, as well as across time. Freedman (2007) is a good example of a state-level analysis, using 25 years of data that straddle changes in various state drunk driving, seat belt, and speed limit laws. The data can be found in the file DRIVING.RAW.

Mullahy and Sindelar (1994) used individual-level data matched with state laws and taxes on alcohol to estimate the effects of laws and taxes on the probability of driving drunk.

- 7 Are blacks discriminated against in the lending market? Hunter and Walker (1996) looked at this question; in fact, we used their data in Computer Exercises C.8 in Chapter 7 and C.2 in Chapter 17.
- 8 Is there a marriage premium for professional athletes? Korenman and Neumark (1991) found a significant wage premium for married men after using a variety of econometric methods, but their analysis is limited because they cannot directly observe productivity. (Plus, Korenman and Neumark used men in a variety of occupations.) Professional athletes provide an interesting group in which to study the marriage premium because we can easily collect data on various productivity measures, in addition to salary. The data set NBASAL.RAW, on players in the National Basketball Association (NBA), is one example. For each player, we have information on points scored, rebounds, assists, playing time, and demographics. As in Computer Exercise C.9 in Chapter 6, we can use multiple regression analysis to test whether the productivity measures differ by marital status. We can also use this kind of data to test whether married men are paid more after we account for productivity differences. (For example, NBA owners may think that married men bring stability to the team, or are better for the team image.) For individual sports—such as golf and tennis—annual earnings directly reflect productivity. Such data, along with age and experience, are relatively easy to collect.
- 9 Answer this question: Are cigarette smokers less productive? A variant on this is: Do workers who smoke take more sick days (everything else being equal)? Mullahy and Portney (1990) use individual-level data to evaluate this question. You could use data at, say, the metropolitan level. Something like average productivity in manufacturing can be related to percentage of manufacturing workers who smoke. Other variables, such as average worker education, capital per worker, and size of the city (you can think of more), should be controlled for.
- 10 Do minimum wages alleviate poverty? You can use state or county data to answer this question. The idea is that the minimum wage varies across states because some states have higher minimums than the federal minimum. Further, there are changes over time in the nominal minimum within a state, some due to changes at the federal level and some because of changes at the state level. Neumark and Wascher (1995) used a panel data set on states to estimate the effects of the minimum wage on the employment rates of young workers, as well as on school enrollment rates.
- 11 What factors affect student performance at public schools? It is fairly easy to get school-level or at least district-level data in most states. Does spending per student matter? Do student-teacher ratios have any effects? It is difficult to estimate *ceteris paribus* effects because spending is related to other factors, such as family incomes or poverty rates. The data set MEAP93.RAW, for Michigan high schools, contains a measure of the poverty rates. Another possibility is to use panel data, or at least to control for a previous year's performance measure (such as average test score or percentage of students passing an exam).

You can look at less obvious factors that affect student performance. For example, after controlling for income, does family structure matter? Perhaps families with two parents, but only one working for a wage, have a positive effect on performance. (There could

be at least two channels: parents spend more time with the children, and they might also volunteer at school.) What about the effect of single-parent households, controlling for income and other factors? You can merge census data for one or two years with school district data.

Do public schools with more charter or private schools nearby better educate their students because of competition? There is a tricky simultaneity issue here because private schools are probably located in areas where the public schools are already poor. Hoxby (1994) used an instrumental variables approach, where population proportions of various religions were IVs for the number of private schools.

Rouse (1998) studied a different question: Did students who were able to attend a private school due to the Milwaukee voucher program perform better than those who did not? She used panel data and was able to control for an unobserved student effect. A subset of Rouse's data is contained in the file `VOUCHER.RAW`.

- 12 Can excess returns on a stock, or a stock index, be predicted by the lagged price/dividend ratio? Or by lagged interest rates or weekly monetary policy? It would be interesting to pick a foreign stock index, or one of the less well-known U.S. indexes. Cochrane (1997) provides a nice survey of recent theories and empirical results for explaining excess stock returns.
- 13 Is there racial discrimination in the market for baseball cards? This involves relating the prices of baseball cards to factors that should affect their prices, such as career statistics, whether the player is in the Hall of Fame, and so on. Holding other factors fixed, do cards of black or Hispanic players sell at a discount?
- 14 You can test whether the market for gambling on sports is efficient. For example, does the spread on football or basketball games contain all usable information for picking against the spread? The data set `PNTSPRD.RAW` contains information on men's college basketball games. The outcome variable is binary. Was the spread covered or not? Then, you can try to find information that was known prior to each game's being played in order to predict whether the spread is covered. (Good luck!) A useful website that contains historical spreads and outcomes for college football and men's basketball games is www.goldsheet.com.
- 15 What effect, if any, does success in college athletics have on other aspects of the university (applications, quality of students, quality of nonathletic departments)? McCormick and Tinsley (1987) looked at the effects of athletic success at major colleges on changes in SAT scores of entering freshmen. Timing is important here: presumably, it is recent past success that affects current applications and student quality. One must control for many other factors—such as tuition and measures of school quality—to make the analysis convincing because, without controlling for other factors, there is a negative correlation between academics and athletic performance. A more recent examination of the link between academic and athletic performance is provided by Tucker (2004), who also looks at how alumni contributions are affected by athletic success.

A variant is to match natural rivals in football or men's basketball and to look at differences across schools as a function of which school won the football game or one or more basketball games. `ATHLET1.RAW` and `ATHLET2.RAW` are small data sets that could be expanded and updated.

- 16** Collect murder rates for a sample of counties (say, from the FBI Uniform Crime Reports) for two years. Make the latter year such that economic and demographic variables are easy to obtain from the *County and City Data Book*. You can obtain the total number of people on death row plus executions for intervening years at the county level. If the years are 1990 and 1985, you might estimate

$$mrd rte_{90} = \beta_0 + \beta_1 mrd rte_{85} + \beta_2 executions + other factors,$$

where interest is in the coefficient on *executions*. The lagged murder rate and other factors serve as controls. If more than two years of data are obtained then the panel data methods in Chapters 13 and 14 can be applied.

Other factors may also act as a deterrent to crime. For example, Cloninger (1991) presented a cross-sectional analysis of the effects of lethal police response on crime rates.

As a different twist, what factors affect crime rates on college campuses? Does the fraction of students living in fraternities or sororities have an effect? Does the size of the police force matter, or the kind of policing used? (Be careful about inferring causality here.) Does having an escort program help reduce crime? What about crime rates in nearby communities? Recently, colleges and universities have been required to report crime statistics; in previous years, reporting was voluntary.

- 17** What factors affect manufacturing productivity at the state level? In addition to levels of capital and worker education, you could look at degree of unionization. A panel data analysis would be most convincing here, using multiple years of census data, say 1980, 1990, 2000, and 2010. Clark (1984) provides an analysis of how unionization affects firm performance and productivity. What other variables might explain productivity?

Firm-level data can be obtained from *Compustat*. For example, other factors being fixed, do changes in unionization affect stock price of a firm?

- 18** Use state- or county-level data or, if possible, school district-level data to look at the factors that affect education spending per pupil. An interesting question is: Other things being equal (such as income and education levels of residents), do districts with a larger percentage of elderly people spend less on schools? Census data can be matched with school district spending data to obtain a very large cross section. The U.S. Department of Education compiles such data.
- 19** What are the effects of state regulations, such as motorcycle helmet laws, on motorcycle fatalities? Or do differences in boating laws—such as minimum operating age—help to explain boating accident rates? The U.S. Department of Transportation compiles such information. This can be merged with data from the *Statistical Abstract of the United States*. A panel data analysis seems to be warranted here.
- 20** What factors affect output growth? Two factors of interest are inflation and investment [for example, Blomström, Lipsey, and Zejan (1996)]. You might use time series data on a country you find interesting. Or you could use a cross section of countries, as in De Long and Summers (1991). Friedman and Kuttner (1992) found evidence that, at least in the 1980s, the spread between the commercial paper rate and the Treasury bill rate affects real output.
- 21** What is the behavior of mergers in the U.S. economy (or some other economy)? Shughart and Tollison (1984) characterize (the log of) annual mergers in the U.S. economy as a

random walk by showing that the difference in logs—roughly, the growth rate—is unpredictable given past growth rates. Does this still hold? Does it hold across various industries? What past measures of economic activity can be used to forecast mergers?

- 22 What factors might explain racial and gender differences in employment and wages? For example, Holzer (1991) reviewed the evidence on the “spatial mismatch hypothesis” to explain differences in employment rates between blacks and whites. Korenman and Neumark (1992) examined the effects of childbearing on women’s wages, while Hersch and Stratton (1997) looked at the effects of household responsibilities on men’s and women’s wages.
- 23 Obtain monthly or quarterly data on teenage employment rates, the minimum wage, and factors that affect teen employment to estimate the effects of the minimum wage on teen employment. Solon (1985) used quarterly U.S. data, while Castillo-Freeman and Freeman (1992) used annual data on Puerto Rico. It might be informative to analyze time series data on a low-wage state in the United States—where changes in the minimum wage are likely to have the largest effect.
- 24 At the city level, estimate a time series model for crime. An example is Cloninger and Sartorius (1979). As a twist, you might estimate the effects of community policing or midnight basketball programs, relatively new innovations in fighting crime. Inferring causality is tricky. Including a lagged dependent variable might be helpful. Because you are using time series data, you should be aware of the spurious regression problem.

Grogger (1990) used data on daily homicide counts to estimate the deterrent effects of capital punishment. Might there be other factors—such as news on lethal response by police—that have an effect on daily crime counts?
- 25 Are there aggregate productivity effects of computer usage? You would need to obtain time series data, perhaps at the national level, on productivity, percentage of employees using computers, and other factors. What about spending (probably as a fraction of total sales) on research and development? What sociological factors (for example, alcohol usage or divorce rates) might affect productivity?
- 26 What factors affect chief executive officer salaries? The files CEOSAL1.RAW and CEOSAL2.RAW are data sets that have various firm performance measures as well as information such as tenure and education. You can certainly update these data files and look for other interesting factors. Rose and Shepard (1997) considered firm diversification as one important determinant of CEO compensation.
- 27 Do differences in tax codes across states affect the amount of foreign direct investment? Hines (1996) studied the effects of state corporate taxes, along with the ability to apply foreign tax credits, on investment from outside the United States.
- 28 What factors affect election outcomes? Does spending matter? Do votes on specific issues matter? Does the state of the local economy matter? See, for example, Levitt (1994) and the data sets VOTE1.RAW and VOTE2.RAW. Fair (1996) performed a time series analysis of U.S. presidential elections.
- 29 Test whether stores or restaurants practice price discrimination based on race or ethnicity. Graddy (1997) used data on fast-food restaurants in New Jersey and Pennsylvania, along with zip code-level characteristics, to see whether prices vary by characteristics of the local population. She found that prices of standard items, such as sodas, increase when the fraction of black residents increases. (Her data are contained in the file DISCRIM.RAW.)

You can collect similar data in your local area by surveying stores or restaurants for prices of common items and matching those with recent census data. See Graddy's paper for details of her analysis.

- 30 Do your own "audit" study to test for race or gender discrimination in hiring. (One such study is described in Example C.3 of Appendix C.) Have pairs of equally qualified friends, say, one male and one female, apply for job openings in local bars or restaurants. You can provide them with phony résumés that give each the same experience and background, where the only difference is gender (or race). Then, you can keep track of who gets the interviews and job offers. Neumark (1996) described one such study conducted in Philadelphia. A variant would be to test whether general physical attractiveness or a specific characteristic, such as being obese or having visible tattoos or body piercings, plays a role in hiring decisions. You would want to use the same gender in the matched pairs, and it may not be easy to get volunteers for such a study.
- 31 Following Hamermesh and Parker (2005), try to establish a link between the physical appearance of college instructors and student evaluations. This can be done on campus via a survey. Somewhat crude data can be obtained from websites that allow students to rank their professors and provide some information about appearance. Ideally, though, any evaluations of attractiveness are not done by current or former students, as those evaluations can be influenced by the grade received.
- 32 Use panel data to study the effects of various economic policies on regional economic growth. Studying the effects of taxes and spending is natural, but other policies may be of interest. For example, Craig, Jackson, and Thomson (2007) study the effects of Small Business Association Loan Guarantee programs on per capita income growth.

List of Journals

The following is a partial list of popular journals containing empirical research in business, economics, and other social sciences. A complete list of journals can be found on the Internet at <http://www.econlit.org>.

American Economic Review
American Journal of Agricultural Economics
American Political Science Review
Applied Economics
Brookings Papers on Economic Activity
Canadian Journal of Economics
Demography
Economic Development and Cultural Change
Economic Inquiry
Economica
Economics Letters
Empirical Economics
Federal Reserve Bulletin
International Economic Review
International Tax and Public Finance

Journal of Applied Econometrics
Journal of Business and Economic Statistics
Journal of Development Economics
Journal of Economic Education
Journal of Empirical Finance
Journal of Environmental Economics and Management
Journal of Finance
Journal of Health Economics
Journal of Human Resources
Journal of Industrial Economics
Journal of International Economics
Journal of Labor Economics
Journal of Monetary Economics
Journal of Money, Credit and Banking
Journal of Political Economy
Journal of Public Economics
Journal of Quantitative Criminology
Journal of Urban Economics
National Bureau of Economic Research Working Papers Series
National Tax Journal
Public Finance Quarterly
Quarterly Journal of Economics
Regional Science & Urban Economics
Review of Economic Studies
Review of Economics and Statistics

Data Sources

Numerous data sources are available throughout the world. Governments of most countries compile a wealth of data; some general and easily accessible data sources for the United States, such as the *Economic Report of the President*, the *Statistical Abstract of the United States*, and the *County and City Data Book*, have already been mentioned. International financial data on many countries are published annually in *International Financial Statistics*. Various magazines, like *BusinessWeek* and *U.S. News and World Report*, often publish statistics—such as CEO salaries and firm performance, or ranking of academic programs—that are novel and can be used in an econometric analysis.

Rather than attempting to provide a list here, we instead give some Internet addresses that are comprehensive sources for economists. A very useful site for economists, called Resources for Economists on the Internet, is maintained by Bill Goffe at SUNY, Oswego. The address is

<http://www.rfe.org>.

This site provides links to journals, data sources, and lists of professional and academic economists. It is quite simple to use.

Another very useful site is

<http://econometriclinks.com>,

which contains links to lots of data sources as well as to other sites of interest to empirical economists.

In addition, the *Journal of Applied Econometrics* and the *Journal of Business and Economic Statistics* have data archives that contain data sets used in most papers published in the journals over the past several years. If you find a data set that interests you, this is a good way to go, as much of the cleaning and formatting of the data have already been done. The downside is that some of these data sets are used in econometric analyses that are more advanced than we have learned about in this text. On the other hand, it is often useful to estimate simpler models using standard econometric methods for comparison.

Many universities, such as the University of California–Berkeley, the University of Michigan, and the University of Maryland, maintain very extensive data sets as well as links to a variety of data sets. Your own library possibly contains an extensive set of links to databases in business, economics, and the other social sciences. The regional Federal Reserve banks, such as the one in St. Louis, manage a variety of data. The National Bureau of Economic Research posts data sets used by some of its researchers. State and federal governments now publish a wealth of data that can be accessed via the Internet. Census data are publicly available from the U.S. Census Bureau. (Two useful publications are the *Economic Census*, published in years ending with two and seven, and the *Census of Population and Housing*, published at the beginning of each decade.) Other agencies, such as the U.S. Department of Justice, also make data available to the public.