Econometrics 120B
Lesson 1:
Why Econometrics?

1. Why study Econometrics?
2. Who needs data anyway?
3. If you had some, what would you do with it?
   Coffee example [Quantifying uncertainty]
   GRE example [Correlation and causality]
4. Types of data:
   Experimental vs. nonexperimental data.
   Cross-sections, Time-Series, Panels.
5. Syllabus & logistics

econ.ucsd.edu/~elib/120b/

© Eli Berman, 2008
1. Why Study Econometrics?

• How do you decide how much to charge for ice cream, ..for your professional services?
• How does Ben Bernanke at the Fed decide how to set interest rates?
• Will theory alone provide an answer?
Example: Demand for Coffee

• Theory tells us only that it's (almost always) downward sloping.

\[ q = \alpha + \beta p \text{ where } \beta < 0 \], assuming linear demand

If we could experiment by offering different prices, how much would individuals purchase at each price?

• Note: 1. Theory doesn't predict the shape of the curve, or the size of \( \beta \). Those must be estimated.
  2. You don't often get to run this experiment.

GOOD DATA IS EXPENSIVE
2. Who needs data anyway?

- Economics as an empirical science. If it can't be used to predict behavior, what would you do with it?
- The cost of data is constantly falling. More data is available for free now than was ever available before.
  - A great opportunity. What would you do with it all?
3. If you had some data, what would you do with it?

- **Econometrics**: The practice of combining economic theory with data to make statistical inferences and predictions

- What do we do with it?
  - write down models,
  - estimate parameters,
  - use those estimated parameters to
    a) predict outcomes and b) test theories
Back to the Coffee example.

- The model: linear demand for coffee

- Basically, we will study how to draw a line through a set of points.
  - the line is our model of coffee demand
  - the points reflect the actual behavior of individuals

- Our set of points is just a sample, and we want to know \( \alpha \) and \( \beta \), which are parameters in some larger population.
Quantifying Uncertainty

• Use statistical theory to assign a standard error to the parameter estimates, $a$ and $b$;
• Interpret those estimates as random variables in a probability distribution.
• Draw confidence intervals for $\alpha$ and $\beta$ around $a$ and $b$ (analogous to confidence intervals for $\mu$, the population mean)
• Then sell the results to the coffee industry, so that profits can be maximized.
Example #2: GRE prep courses

- How much does a GRE prep course increase your GRE score?
- How might GRE scores vary by preparation?

\[ Y = \alpha + \beta x + \varepsilon \]

Causal solutions: 1) Before-After
2) Randomly Assign
3) Include other into 0

\[ \text{Causal Score} \]

- Don't
- Take Course
- True
- Causal Scope

1 \rightarrow X
Correlation vs. Causality

• Does the negative correlation imply that prep courses lower your GRE score?

• Omitted variable bias:
  \[ y = \alpha + \beta x_1 + \gamma x_2 \], where \( x_2 \) is omitted GRE ability

• Why didn’t this problem show up in the coffee example?
  - experimental vs. nonexperimental data
4. Getting Good Data

- Experimental data vs. nonexperimental
  - preferably with random assignment of “x”
  - why is that important? Think of nonexperimental data on coffee sales.

- Cross-Sections: many units in one period
- Time-Series: one unit over many periods
- Panels: many units over many periods
<table>
<thead>
<tr>
<th>Observation (district) number</th>
<th>District average test score (fifth grade)</th>
<th>Student-teacher ratio</th>
<th>Expenditure per pupil ($)</th>
<th>% of students learning English</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>690.8</td>
<td>17.89</td>
<td>$6,385</td>
<td>0.0%</td>
</tr>
<tr>
<td>2</td>
<td>661.2</td>
<td>21.52</td>
<td>5,099</td>
<td>4.6</td>
</tr>
<tr>
<td>3</td>
<td>643.6</td>
<td>18.70</td>
<td>5,502</td>
<td>30.0</td>
</tr>
<tr>
<td>4</td>
<td>647.7</td>
<td>17.36</td>
<td>7,102</td>
<td>0.0</td>
</tr>
<tr>
<td>5</td>
<td>640.8</td>
<td>18.67</td>
<td>5,236</td>
<td>13.9</td>
</tr>
<tr>
<td>418</td>
<td></td>
<td></td>
<td>4,403</td>
<td>24.3</td>
</tr>
<tr>
<td>419</td>
<td></td>
<td></td>
<td>4,776</td>
<td>3.0</td>
</tr>
<tr>
<td>420</td>
<td></td>
<td></td>
<td>5,993</td>
<td>5.0</td>
</tr>
</tbody>
</table>

*Note:* The California test score data set is described in Appendix 4.1.
<table>
<thead>
<tr>
<th>Observation number</th>
<th>Date (Year:quarter)</th>
<th>CPI inflation rate (% per year at an annual rate)</th>
<th>Unemployment rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1959:II</td>
<td>0.7%</td>
<td>5.1%</td>
</tr>
<tr>
<td>2</td>
<td>1959:III</td>
<td>2.1</td>
<td>5.3</td>
</tr>
<tr>
<td>3</td>
<td>1959:IV</td>
<td>2.4</td>
<td>5.6</td>
</tr>
<tr>
<td>4</td>
<td>1960:I</td>
<td>0.4</td>
<td>5.1</td>
</tr>
<tr>
<td>5</td>
<td>1960:II</td>
<td>2.4</td>
<td>5.2</td>
</tr>
<tr>
<td>165</td>
<td>2000:II</td>
<td>3.0</td>
<td>4.0</td>
</tr>
<tr>
<td>166</td>
<td>2000:III</td>
<td>3.5</td>
<td>4.0</td>
</tr>
<tr>
<td>167</td>
<td>2000:IV</td>
<td>2.8</td>
<td>4.0</td>
</tr>
</tbody>
</table>

*Note:* The U.S. inflation and unemployment data set is described in Appendix 12.1.
TABLE 1.3  Selected Observations on Cigarette Sales, Prices, and Taxes, by State and Year for U.S. States, 1985–1995

<table>
<thead>
<tr>
<th>Observation number</th>
<th>State</th>
<th>Year</th>
<th>Cigarette sales (packs per capita)</th>
<th>Average price per pack (including taxes)</th>
<th>Total taxes (cigarette excise tax + sales tax)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Alabama</td>
<td>1985</td>
<td>116.5</td>
<td>$1.022</td>
<td>$0.333</td>
</tr>
<tr>
<td>2</td>
<td>Arkansas</td>
<td>1985</td>
<td>128.5</td>
<td>1.015</td>
<td>0.370</td>
</tr>
<tr>
<td>3</td>
<td>Arizona</td>
<td>1985</td>
<td>104.5</td>
<td>1.086</td>
<td>0.362</td>
</tr>
<tr>
<td>47</td>
<td>West Virginia</td>
<td>1985</td>
<td>112.8</td>
<td>1.089</td>
<td>0.382</td>
</tr>
<tr>
<td>48</td>
<td>Wyoming</td>
<td>1985</td>
<td>129.4</td>
<td>0.935</td>
<td>0.240</td>
</tr>
<tr>
<td>49</td>
<td>Alabama</td>
<td>1986</td>
<td>117.2</td>
<td>1.080</td>
<td>0.334</td>
</tr>
<tr>
<td>96</td>
<td>Wyoming</td>
<td>1986</td>
<td>127.8</td>
<td>1.007</td>
<td>0.240</td>
</tr>
<tr>
<td>97</td>
<td>Alabama</td>
<td>1987</td>
<td>115.8</td>
<td>1.135</td>
<td>0.335</td>
</tr>
<tr>
<td>528</td>
<td>Wyoming</td>
<td>1995</td>
<td>112.2</td>
<td>1.585</td>
<td>0.360</td>
</tr>
</tbody>
</table>

Note: The cigarette consumption data set is described in Appendix 10.1.
5. Syllabus and Logistics

UC San Diego
Winter 2008
Class: T-Th 11-12:20 (Center 115)

Professor Eli Berman
elib@ucsd.edu
OH: T 9:30-10:30 EC 218

Economics 120B
Econometrics

This course prepares students for practical empirical research in an academic or business setting. It introduces three major ideas in econometrics: quantifying uncertainty using confidence intervals, using regression to infer causal relationships, and using regression for prediction. We also cover exotic concepts such as heteroskedasticity and instrumental variables. Graduates will be able to conduct and read simple empirical research.

The prerequisite is one of EC 120A, Math 183, or ECE109. Students without a prerequisite are welcome to attend the first few classes but must get permission to register.

The material is fairly well covered in Introduction to Econometrics (first or second edition), by Jim Stock and Mark Watson, which is required reading and is available at the bookstore. Another helpful text is Introductory Econometrics, by Arthur Goldberger.

We will learn to use an econometrics program called Stata. Students have access to Stata in the computer lab. Individual copies of Small Stata can be leased for about $48 at http://www.stata.com/order/new/edu/gradplans/gp-direct.html. Stata is essential for problem sets.
Evaluation, etc..

Evaluation: There will be a midterm exam (worth about 35% of the grade) on Thursday February 7, a final (about 45%) on Thursday, March 20 at 11:30 AM, and three or four problem sets (worth about 20% together). No other exams are planned. Exams may be proctored using video. Reviewing class notes with knowledgeable friends, solving problem sets, reviewing old exams, and reading the text are good ways to prepare for exams.

Written work: The “Buckley” waiver makes it much easier to return your written work, by allowing us to put it in a public place rather than having you meet with the TA. Please read the attached waiver, sign it, and return it to a TA if you’re comfortable with it.

Disabilities will be accommodated. Please refer to OSD policy on what we do to help.

If you have any questions please feel welcome to come chat in office hours. Office hours tend to get congested the week before exams.

In order to compensate for lost classes in 120A during the week of wildfires we may shift review classes in preparation for the midterm and final exam to the evening.
COURSE OUTLINE:

1. Introduction: Why Study Econometrics?

2. Probability and Statistics: A quick review
Probability, random variables, the normal distribution and the central limit theorem, inference, confidence intervals and hypothesis testing. Asymptotics of the sample mean. Using Stata. Reading: Chapters #2 and #3.

3. Simple Regression (one regressor)
Fitting a line through a cloud of points. Least squares, unbiased estimates, consistent estimates, confidence intervals, hypothesis testing, omitted variable bias, $R^2$. Reading: Chapters #4 and #5.

{Review and midterm about here}

4. Multiple Regression: Estimation
The second explanatory variable, interpreting coefficients, efficiency & heteroskedasticity, omitted variable bias. Reading: Chapter #6.
5. Causal Inference and Random Assignment
Random assignment vs. omitted variable bias. Reading: Ch #13.

6. Multiple Regression: Inference
Confidence intervals (CI) for parameters, CI for predictions, hypothesis testing, single (t) vs. multiple (F) tests. Etiquette for reporting results.
Reading: Chapter #7.

7. Sources of Bias: measurement error, sample selection, simultaneity and omitted variables
Omitted Variable Bias again, Measurement Error, Fixed Effects, Sample Selection, Simultaneity.
Reading: Chapters #9 and #10.

8. Identification and Instrumental Variables
Causal inference again, instrumental variables vs. omitted variable bias, two-stage least squares, natural experiments.
Readings: Chapter #12;
Summary: Why Econometrics?

1. Why study Econometrics?

2. Who needs data anyway?

3. If you had some, what would you do with it?
   - Coffee example [Quantifying uncertainty]
   - GRE example [Correlation and causality]
   *These two themes will recur throughout the course*

4. Types of data:
   - Experimental vs. nonexperimental data.
   - Cross-sections, Time-Series, Panels.

5. Syllabus & logistics - econ.ucsd.edu/~elib/120b

Next.. Statistics Review

*Don’t forget: Problem Set #1*