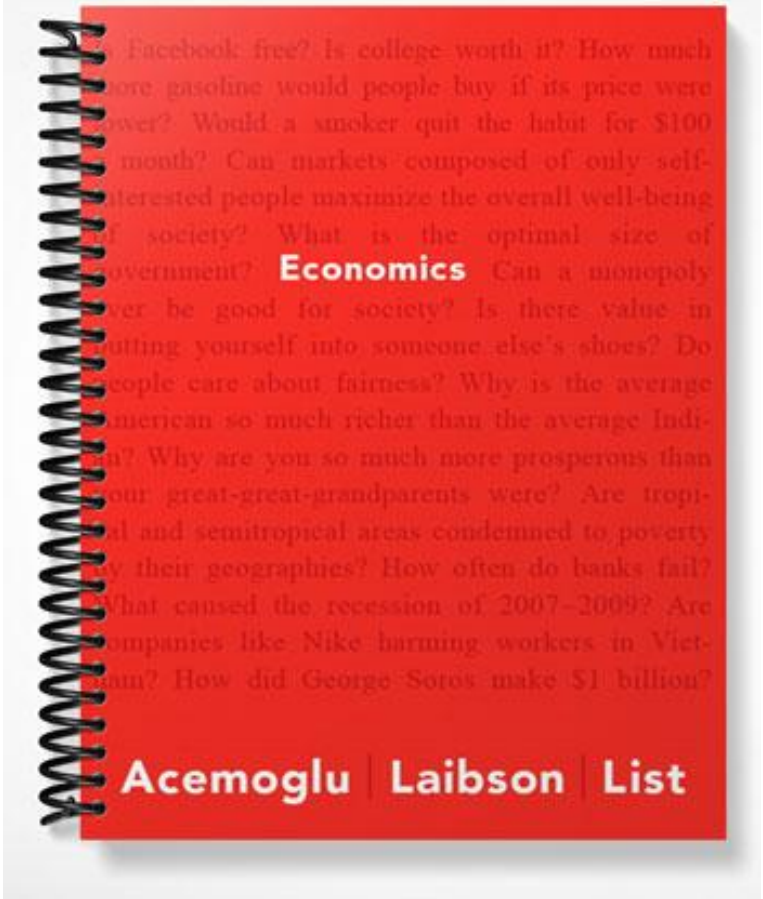
# SOLUTIONS MANUAL



Chapter 2

# Economic Methods and Economic Questions

## Questions

1. What does it mean to say that economists use the scientific method? How do economists distinguish between models that work and those that don't?

   *Answer*: The scientific method is the name for the ongoing process that economists and other scientists use to develop models of the world, test those models with data and evaluate how well they predict or describe behavior. While this process may not reveal the 'true' model of the world, it does help in identifying models that are useful in understanding the world.

   In order to decide whether models make accurate predictions or not, economists test them against real-world data. Data are facts, measurements, or statistics that describe the world. This process of testing models against data is called empiricism.

2. What is meant by empiricism? How do empiricists use hypotheses?

   *Answer*: Empirical evidence is a set of facts established by observation and measurement, which are used to evaluate a model. Empiricism refers to the practice of using data to test economic models. When conducting empirical analysis, economists refer to a model's predictions as hypotheses. Hypotheses are predictions (typically generated by a model) that can be tested with data.

3. What are two important properties of economic models? Models are often simplified descriptions of a real-world phenomenon. Does this mean that they are unrealistic?

   *Answer*: A good economic model has two important properties. First, it is an approximation. The model predicts what would happen on average. Second, it makes predictions that can be tested with data.

   A model is a simplified description, or representation, of reality. Because models are simplified, they are not perfect replicas of reality. However, this does not mean that they are unrealistic. Models are usually simplified in order to be able to isolate the relationship between two variables. Even if a model is based on simplified assumptions, it may still help us make good predictions and plan for the future.

4. How is the mean calculated from a series of observations? Suppose 5,000 people bought popsicles on a hot summer day. If the mean of the average number of popsicles that each person bought is 2, how many popsicles were sold that day?

   *Answer*: The mean is the average value of a set of observations. It is calculated as the sum of all the different items divided by the number of items.

   The average value is the sum of all popsicles sold divided by the number of people who bought them. If each of the 5,000 people bought an average of 2 popsicles, that means that 10,000 popsicles were sold that day.

5. How does the sample size affect the validity of an empirical argument? When is it acceptable to use only one example to disprove a statement?

   *Answer*: The size of the sample used to test the argument can affect the results. A small sample may bias the results of a study. A key strength of economic analysis is the amount of data used. Using a large number of observations strengthens the force of an empirical argument. For example, if you collect information on consumption from 20,000 people as opposed to 20 people, you are likely to get a more representative result. A single example can be used to contradict a statement. For example, a single black swan can disprove the statement that all swans are white.

6. Explain why correlation does not always imply causation. Does causation always imply *positive* correlation? Explain your answer.

   *Answer*: Correlation means that there is a relationship between two variables; as one variable changes, another variable changes. Causation occurs when one variable directly affects another through a cause-and-effect relationship. Correlation suggests that there is some kind of connection, but not necessarily a cause and an effect. For example, the number of storks in a region might be correlated with the number of babies born in the region. But this doesn't mean that storks bring babies.

   Positive correlation implies that two variables tend to move in the same direction. However, causation need not only imply positive correlation. For example, an increase in the price of bacon may cause people to buy smaller amounts of bacon. In this example, the price of bacon and the quantity of bacon purchased will show a negative correlation.

7. Give an example of a pair of variables that have a positive correlation, a pair of variables that have a negative correlation, and a pair of variables that have zero correlation.

   *Answer*: A person's IQ and his or her telephone number are likely to show zero correlation. The number of winter coats sold and the temperature outside are likely to show a negative correlation. The quantity of fertilizers used and crop yield (e.g., the number of bushels of wheat grown per acre) are likely to have a positive correlation.

8. What is meant by randomization? How does randomization affect the results of an experiment?

   *Answer*: Randomization is the assignment of subjects by chance, rather than by choice, to a test group or control group. Assigning participants randomly will ensure that the result of the experiment is not biased. For example, suppose students with poor scores are assigned to one type of teaching method and students with good scores are assigned to another type of teaching method. It will be difficult to decide which teaching method is more effective as the students with the higher scores are likely to do better than the students with the poorer scores, irrespective of the teaching method used.

9. This chapter discussed natural and randomized experiments. How does a natural experiment differ from a randomized one? Which one is likely to yield more accurate results?

   *Answer*: A natural experiment is an empirical study in which some process – out of the control of the experimenter – has assigned subjects to control and test groups in a random or nearly random way. The process of randomization involves the assignment of subjects by chance, rather than by choice, to a test group or control group. The test group and the control group are treated identically, except along a single dimension that is intentionally varied across the two groups. The impact of this variation is the focus of the experiment. Both types of experiments can yield accurate results. Natural experiments are likely to be used when there are budget or time constraints to conducting a randomized experiment.

10. Suppose you had to find the effect of seatbelt rules on road accident fatalities. Would you choose to run a randomized experiment or would it make sense to use natural experiments here? Explain.

*Answer:* It would be difficult (and, in many people's view, unethical) to conduct a randomized experiment. Instead, the study should use a natural experiment. You can study data on the causes of road accident fatalities in cities where seatbelt rules were not enforced, or in cities that have recently adopted new, more stringent seat belt laws. Controlling for other factors like an increase in the number of cars, etc., you can then look at similar data when seatbelt rules have been implemented.

## Problems

1. This chapter talks about means. The median is a closely related concept. The median is the numerical value separating the higher half of your data from the lower half. You can find the median by arranging all of the observations from lowest value to highest value and picking the middle value (assuming you have an odd number of observations). While the mean and median are closely related, the difference between the mean and the median is sometimes of interest.

   a. Suppose country A has five families. Their incomes are $10,000, $20,000, $30,000, $40,000, and $50,000. What is the median family income in A? What is the mean income?

   b. Country B also has five families. Their incomes are $10,000, $20,000, $30,000, $40,000, and $150,000. What is the median family income in B? What is the mean income?

   c. In which country is income inequality greater, A or B?

   d. Suppose you thought income inequality in the US had increased over time. Based on your answers to this question, would you expect that the ratio of the mean income in the US to the median income has risen or fallen? Explain.

   *Answer:*

      a. We can find the mean by summing the observations and dividing by the number of observations. So the mean income in Country A is ($10,000 + $20,000 + $30,000 + $40,000 + $50,000) / 5 = $30,000. The median income is the income of the family in the middle of the income distribution. The median income in Country A is $30,000. Two families have income below $30,000 and two have income above $30,000.

      b. A similar argument shows that the mean income in Country B is ($10,000 + $20,000 + $30,000 + $40,000 + $150,000) / 5 = $50,000. Median income in B is $30,000; as in Country A, two families have income below $30,000 and two have income above $30,000.

      c. Income inequality is higher in Country B. The highest income family in Country B earns $150,000, 60% of the total income in that country. The highest income family in A earns $50,000, just 33% of total income in A. We found that the median income in the two countries was the same but the mean income was very different. Means will be heavily influenced by extreme values such as the incomes of the very wealthy; median income is less sensitive to extremes. Economists sometimes use the ratio of the mean to median income in a country as a rough measure of income inequality; higher values of this ratio reflect greater inequality.

      d. You should expect to find that the ratio of the mean to median income has risen. As we argued above, the mean is more sensitive than the median to the incomes of the very wealthy.

2. Consider the following situation: your math professor tells your class that the mean score on the final exam is 43. The exam was scored on a total of 100 points. Does this imply that you, too, scored poorly on the exam?

   *Answer*: The mean is an average value of a set of observations. Although the mean score is representative of the values in the set of observations, a single value (for example, your exam score)

could be very different from the mean. These 'outliers' could skew the mean toward being lower or higher than most of the other observations. So, it could be possible that you scored higher than, say 75, but your classmates' lower scores pulled the value of the mean downward.

3. This chapter stressed the importance of using appropriate samples for empirical studies. Consider the following two problems in that light.

   a. You are given a class assignment to find out if people's political leanings affect the newspaper or magazine that they read. You survey two students taking a political science class and five people at a coffee shop. Almost all the people you have spoken to tell you that their political affiliations do not affect what they read. Based on the results of your study, you conclude that there is no relationship between political inclinations and the choice of a newspaper. Is this a valid conclusion? Why or why not?

   b. Your uncle tells you that the newspaper or magazine that people buy will depend on their age. He says that he believes this because, at home, his wife and his teenage kids read different papers. Do you think his conclusion is justified?

   *Answer:*

      a. The conclusion is not likely to be valid as the sample used in the study is too small. Convincing data analysis will depend on a much larger sample of people. The subjects of the study should also be randomly chosen to minimize the possibility of the results being biased.

      b. This is an example of argument by anecdote. Using a small sample of people to judge a statistical relationship is likely to lead to flawed conclusions. The fact that your uncle's wife and kids do not base their reading on their political affiliations does not mean that others do not. In order to arrive at a conclusion, you need to survey more people and also make sure that they are chosen randomly.

4. Some studies have found that people who owned guns were more likely to be killed with a gun. Do you think this study is strong evidence in favor of stricter gun control laws? Explain.

   *Answer*: Not necessarily. It is quite possible that people who thought they were at risk (perhaps because they live in dangerous neighborhoods) were more likely to buy a gun for self-protection. This is an example of a case where correlation may not imply causation. There has been a good deal of research on this question. See, for example, a Harvard School of Public Health 2011 interview with David Hemenway (http://www.hsph.harvard.edu/news/features/review-guns-politics-hemenway).

5. As the text explains, it can sometimes be very difficult to sort out the direction of causality.

   a. Why might you think more police officers would lead to lower crime rates? Why might you think that higher crime rates would lead to more police officers?

   b. In 2012, the *New England Journal of Medicine* published research that showed a strong correlation between the consumption of chocolate in a country and the number of Nobel Prize winners in that country. Do you think countries that want to encourage their citizens to win Nobel Prizes should increase their consumption of chocolate?

   *Answer:*

      a. There is a great deal of evidence that increasing the number of police officers in a neighborhood can drive down crime. The police, for example, will deter criminals who realize the chances they will be caught have gone up and the police may be able to head off conflicts between gangs. Therefore more police could lead to less crime. Cities strategically assign more police to high crime areas (since by definition, those are the areas where crimes are more likely to occur). Therefore, more crime can lead to more police.

b. Correlation does not necessarily imply causation. A strong positive correlation between chocolate consumption and Nobel Prize winners does not, by itself, suggest causation. It is possible that this is a chance correlation. It may also be the case that certain variables that could explain this relationship have been omitted from the study.

See the June 11, 2013 New York Times article "Chicago Tactics Put Major Dent in Killing Trend" (http://www.nytimes.com/2013/06/11/us/chicago-homicides-fall-by-34-percent-so-far-this-year.html?hp) on the relationship between police and crime rates. See http://www.reuters.com/article/2012/10/10/us-eat-chocolate-win-the-nobel-prize-idUSBRE8991MS20121010 on the effects of eating chocolate.

6. The chapter shows that as a general rule people with more education earn higher salaries. Economists have offered two explanations of this relationship. The human capital argument says that high schools and colleges teach people valuable skills and employers are willing to pay higher salaries to attract people with those skills. The signaling argument says that college graduates earn more because a college degree is a signal to employers that a job applicant is diligent, intelligent, and persevering. How might you use data on people with two, three, and four years of college education to shed light on this controversy?

   *Answer*: If the human capital explanation is correct, then we might expect to find that people who attend college but do not graduate earn salaries that are close to what college graduates earn. Consider the extreme case of people who drop out of college the week before graduation. It is very unlikely that they would have improved their job skills much in that last week. The human capital school of thought would suggest that they should therefore earn roughly the same salaries as college graduates. On the other hand, the signaling school of thought would argue that these people should earn significantly less than college graduates. Employers would interpret their failure to graduate as a signal they are not as diligent or persevering as people who see their college educations through to the end. There is substantial literature on what is often called the "sheepskin effect" (college diplomas used to be written on sheepskin; Notre Dame continued to use sheepskin until 2012). That literature suggests that human capital and signaling both contribute to the returns to education that we observe in the data.

   For recent evidence, see the Michael Greenstone and Adam Looney 2013 Brookings Institution study **"**Is Starting College and Not Finishing Really That Bad?" (http://www.brookings.edu/blogs/jobs/posts/2013/06/07-return-to-some-college-greenstone-looney). They find that people with some college education but who do not graduate from college earn an average of $8,000 more per year than high school graduates who never attend college.

7. Maimonides, a twelfth century scholar, said, "Twenty-five children may be put in the charge of one teacher. If the number in the class exceeds twenty-five but is not more than forty, he should have an assistant to help with the instruction. If there are more than forty, two teachers must be appointed." Israel follows Maimonides's rule in determining the number of teachers for each class. How could you use Maimonides's rule as a natural experiment to study the effect of teacher-student ratios on student achievement?

   *Answer*: Maimonides's rule generates a natural experiment to study of the effect of class size. Suppose School A has 40 third-graders and School B has 41. Following Maimonides's rule, School A would have one 40-student class while School B would have one 20-student class and one 21-student class. Everything else equal, if smaller classes improve achievement then we should expect to see higher test scores in School B.

   See Joshua D. Angrist and Victor Lavy "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics*, 1999 (http://qje.oxfordjournals.org/content/114/2/533.full.pdf+html).

8. Oregon expanded its Medicaid coverage in 2008. Roughly 90,000 people applied but the state had funds to cover only an additional 30,000 people (who were randomly chosen from the total applicant pool of 90,000). How could you use the Oregon experience to estimate the impact of increased access to health care on health outcomes?

   *Answer*: The Oregon experience is a natural experiment. The state chose people randomly from the pool of applicants, and so on average the new Medicaid recipients were very similar to the people who applied but were turned down. By tracking the health outcomes of people in these two groups we can study the effect of better access to health care.

   See "Medicaid Access Increases Use of Care, Study Finds," New York Times, May 1, 2013 (http://www.nytimes.com/2013/05/02/business/study-finds-health-care-use-rises-with-expanded-medicaid.html?_r=0).

9. A simple economic model predicts that a fall in the price of bus tickets means that more people will take the bus. However, you observe that some people still do not take the bus even after the price of a ticket fell.

   a. Is the model incorrect?

   b. How would you test this model?

   *Answer:*

   a. The model is not incorrect. Models are only approximations of real-life behavior. Even very good models make predictions that are often correct. So, on average, more people will take the bus. The model is also likely to have made some assumptions, such as no change in costs of other types of transport, or that people have no specific preferences and cost is the only determinant of the mode of transport used. In reality, some of these assumptions may be violated which could explain why a fall in the price of bus tickets does not induce everyone to take the bus. That does not imply that the model's conclusion is incorrect. In situations where the assumptions it makes are satisfied, its prediction will often be correct.

   b. The hypothesis here states that as bus prices fall, the number of passengers who take the bus will increase. A natural experiment can be used to test this model. You can use data on price changes and changes in revenues earned from tickets to see whether the model is accurate.

A1. How would you represent the following graphically?

   a. Income inequality in the U.S. has increased over the past 10 years.

   b. All the workers in the manufacturing sector in a particular country fit into one (and only one) of the following three categories: 31.5 percent are high school dropouts, 63.5 percent have a high school diploma, and the rest have vocational training certificates.

   c. The median income of a household in Alabama was $43,464 in 2012, and the median income of a household in Connecticut was $64,247 in 2012.

   *Answer:*

   a. Since the graph needs to show how income inequality increases over a period of time, a time-series graph needs to be used here.

   b. A pie chart is a circular chart split into segments to show the percentages of parts to the whole. Since the given data is in percentages, a pie-chart can be used to represent each category of workers.

   c. A bar chart would be a good way to compare income in Alabama and Connecticut. The height of each bar would represent the income in each one of the states.

A2. Consider the following data that shows the quantity of coffee produced in Brazil from 2004-2012.

| Year | Production (in tons) |
|---|---|
| 2004 | 2,465,710 |
| 2005 | 2,140,169 |
| 2006 | 2,573,368 |
| 2007 | 2,249,011 |
| 2008 | 2,796,927 |
| 2009 | 2,440,056 |
| 2010 | 2,907,265 |
| 2011 | 2,700,440 |
| 2012 | 3,037,534 |

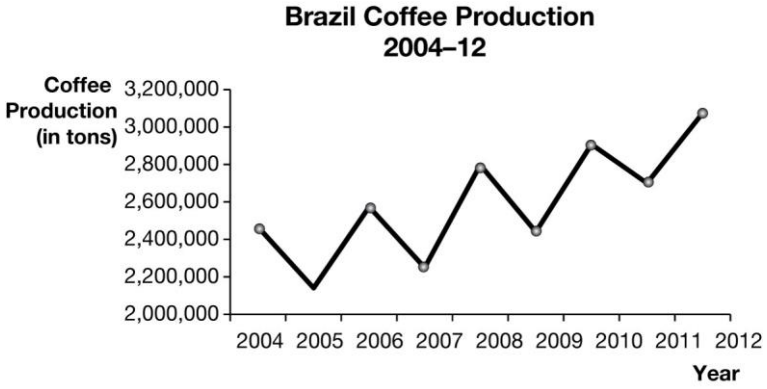
a. Plot the data in a time series graph.

b. What is the mean quantity of coffee that Brazil produced from 2009 to 2011?

c. In percentage terms, how much has the 2012 crop increased over the 2009-2011 mean?

*Answer:*

a. A time-series graph can be used to represent the quantity of coffee produced from 2004 to 2012.




b. The average quantity of coffee that Brazil produced during the 2009-11 period is 2,682,587 tons. This is the sum of the total quantity produced divided by the number of years.

c. The coffee crop in 2012 is 14.6% larger than the average coffee crop in 2009-2011. The increase in production is 3,037,534 – 2,682,587 = 354,947. In percentage terms, the change is 354,947/ 2,682,587 = 13.2%.

Data taken from: http://faostat.fao.org/site/567/default.aspx#ancor

A3. Suppose the following table shows the relationship between revenue that the Girl Scouts earn and the number of cookie boxes that they sell.

| Number of cookie boxes | Revenue ($) |
|---|---|
| 50 | 200 |
| 150 | 600 |
| 250 | 1,000 |
| 350 | 1,400 |
| 450 | 1,800 |
| 550 | 2,200 |

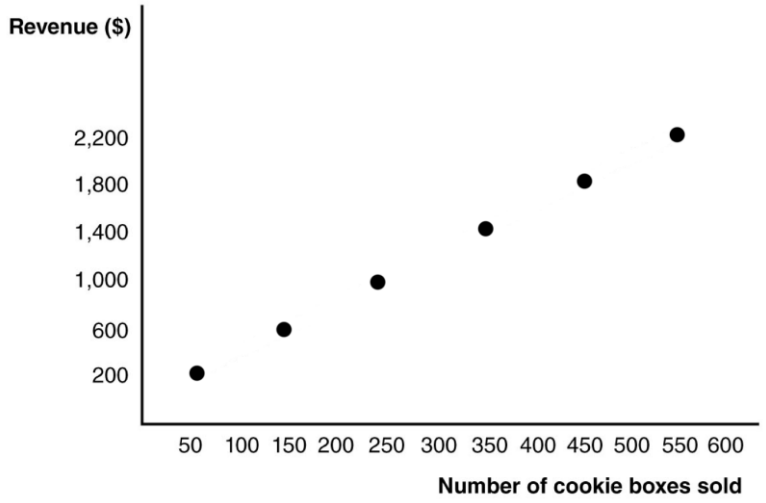
a. Present the data in a scatter plot.

b. Do the two variables have a positive relationship or do they have a negative relationship? Explain.

c. What is the slope of the line that you get in the scatter plot? What does the slope imply about the price of a box of Girl Scout cookies?

*Answer:*

a. The following line chart shows the relationship between the Girl Scouts' revenue and the number of cookie boxes that they sell:




b. Since the values of both variables increase together in the same direction, they have a positive relationship. This means that as more cookie boxes are sold, the revenue earned increases.

c. The slope is constant in this problem and so we can choose any two points to calculate the slope. Suppose we use the first and last data points. The slope is calculated as $\frac{\text{Change in revenue}}{\text{Change in cookie boxes sold}} = \frac{2200-200}{550-50} = \frac{2000}{500} = 4$. The slope implies that one extra box of cookies sold is associated with $4 more in revenue.