**[MODULE I  
STATISTICS AND TIME SERIES](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html" \l "pt1)**

[**CHAPTER 1   
Introduction and Context**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1)

[**1.1 What Is Forecasting?**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.1)

Most people have an intuitive notion of what forecasting means. In our daily lives, we refer very frequently to *future* events, we look *forward*, we have the *foresight* to do something, we are able to *foretell*, we *foresee* an event, and we say that something is *forthcoming*. Forecasting, implicit or explicitly, is embedded in all our planning activities from the beginning of our history.

[**1.1.1 The First Forecaster in History: The Delphi Oracle**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.1.1)

The first forecaster recorded in history was in the city of Delphi in ancient Greece (6th century BC until 2nd century AD) known as the **Delphi Oracle**[1](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#fn1_1) to which pilgrims from all over the world came to seek advice from the oracle in the temple of Apollo. This is a brief description of how this forecaster worked and one of its most famous predictions.

The pilgrims submitted questions to the oracle through emissaries, and every month, Pythia, the priestess of the god Apollo, spoke the prophecies (in our jargon, conveyed fore- casts). According to history, Pythia sat on a tripod chewing bay leaves, drinking water from the Kassotis spring, and inhaling the fumes that sent her into a trance. Then she spoke the forecast, which purposely was ambiguous and subject to interpretation because the Gods cannot be wrong! (Perhaps there is some lesson here for professional forecasters.)

A famous oracle came around 480 BC during the Persian wars when the Athenians were fighting the Persians, who had an enormous float of wooden boats and a large army of men. Emissaries were sent to the Oracle of Delphi to solicit a prediction on the outcome of the war. The Oracle spoke these words “Only wooden walls will save Athens.” Obviously, this forecast needed interpretation. Thank god for the wisdom of Themistocles, who understood the prediction: If the Athenians could construct flexible small fast boats that fit the geographical configuration of the gulf of Athens, they would outmaneuver Persia’s large and slow boats and win the war.

This is an example of a historical *forecast*, which required some skills to be under-stood. Not much has changed in modern times. We do not have a physical place to go to listen to the oracle, but we have an army of professional forecasters in the public and private sectors whose job is to foretell the future.

[**1.1.2 Examples of Modern Forecasts**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.1.2)

All socioeconomic issues of the day require some level of prediction. Open the newspaper and just read some headlines: the exit strategy of the Federal Reserve, health care reform, financial regulation, systemic risk, network traffic, the fear of inflation, job creation, and so on. The following are some examples of how the daily business press presents forecasting. As you read them, think about three questions:

**1.** What is the predicted event?

**2.** What is the magnitude of the predicted event?

**3.** What is the future date of the predicted event?

•*On the bond and stock markets:* On March 22, 2010, a *Fortune* magazine journalist asked Wilbur Ross, an American investor, how he saw 2010. Ross answered that the year would be volatile and that some stocks would perform well despite the environment.

•*In U.S. industrial production:* On March 16, 2010, *The New York Times* published an article claiming that production would rise in 2010 citing, among others, John Ryding, chief economist for RDQ Economics, who forecast 7% growth in industrial production in 2010 as a consequence of improvements in the labor market and more spending on capital goods in foreign countries.

•*On network traffic:* On February 27, 2010, *The Economist* published a special report on managing information stating that, according to *Cisco*, annual Internet traffic by 2013 would reach 667 exabytes (one exactabyte equals one billion gigabites).

•*On U.S. bank bailouts:* On March 22, 2010, *Fortune* magazine published an article on the state of bank bailouts, citing a prediction by the U.S. Treasury Department that the cost of the Troubled Asset Relief Program (TARP) to taxpayers would eventually be $117 billion (adjusted for inflation).

•*On U.S. real estate:* On April 2010, *Kiplinger’s* magazine ran an article on the state of the rental real estate market in which experts expected that the national vacancy rate would be 7.8% by the end of 2010, would rapidly recover beginning in 2011, and would experience strong rent growth from 2011 to 2015.

•*On India’s economic growth: The Economist* published a report on India on March 2010 stating that the Indian economy would grow by 7.2% by the end of the 2010 fiscal year (March 31) with the expectation of 9% growth rate in the medium term.

•*On small business:* On March 15, 2010, *The Wall Street Journal* published predictions by Raj Date who expected a shortfall of $250 to $500 billion in lending to small business when the economy begins to recover.

What we gather from these examples is that some predictions are detailed and some are sketchy concerning either the magnitude of the future event or its date of occurrence. Examples 2, 3, 5, and 6 are very precise on the magnitude and timing of the prediction: 7% growth in industrial production in 2010; 667 exabytes by 2013; 7.8% vacancy rate by year-end 2010; and 7.2% growth in the Indian economy by March 31, 2010. Examples 1, 4, and 7 are fuzzier predictions: 2010 would be a volatile year (how much volatility?), TARP eventually (when?) would cost taxpayers $117 billion, and lending shortfall for small business would be as much as $250 billion to $500 billion as the economy recovered (when?). Furthermore, a common feature to these examples is that we, the readers, do not have any information about how certain or uncertain the forecaster is about the predictions. Only when time passes can we judge who was right and who was wrong.

These introductory examples show the relevant features to answer the question, “What is forecasting?”

[**1.1.3 Definition of Forecasting**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.1.3)

We define **forecasting** as the *science* and the *art* to predict a *future* event with some degree of *accuracy*.

**1.1.3.1 Why Is Forecasting a Science?**

You may be wondering how it is possible to state a precise statement such as “we expect 7% growth in U.S. industrial production in 2010.” To produce such statements, we need a methodological approach to summarize and analyze the information that is available. Statistical and mathematical methods are very useful in discovering time patterns in historical data. We will construct models that synthesize the past and will explain how today is related to yesterday, yesterday to the day before yesterday, and so on. This time dependence is the key to making statements about the future. The mastery of these statistical methods will enable us to construct forecasts that have a logical consistency.

**1.1.3.2 Why Is Forecasting an Art?**

Statistical methods have limitations. They depend on a set of assumptions, which may or may not be satisfied by the available data. Models, by construction, are limited representations of the economic and business environments. In addition, there is also a technological limitation given by the frontier of statistical and mathematical research. Professional forecasters also accumulate soft human capital, that is, knowledge and experiences, which we do not know yet how to quantify or formalize but nevertheless are useful in modifying the forecast provided by a statistical model. In this sense, the forecaster needs judgment and when exercising it, forecasting becomes an art.

**1.1.3.3 Why Do We Care About the Future?**

This seems to be obvious. The past is known to a certain extent, and although statistical models can also be used for backcasting or backtesting, forecasting eventually is always an activity that involves an assessment of future events.

#### 1.1.3.4 Why Does Accuracy Matter?

Professional forecasters succeed by making accurate forecasts. They need to establish a performance record for their activities to be relevant and useful. Broadly speaking, we should not expect a forecast to be exactly accurate in a mathematical sense, that is, hitting the exact future value of the variable of the interest, but we should expect the forecast to be statistically sound and, when possible, also to offer a measure of the uncertainty of the predictions.

### [1.1.4 Two Types of Forecasts](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.1.4)

Customarily, we distinguish between event forecast and time series forecast, although both can be related.

An **event forecast** refers to the future occurrence of an outcome and/or the timing of such an occurrence. For instance, read the following questions and provide your best event forecast:

•Will the Federal Reserve raise interest rates at its next board meeting?

•When will the recession end?

•When will the stock market recover?

•Will the euro keep appreciating?

Some questions will require just “yes” or “no” answers, and some others will require a future date or a timeline.

The term **time series forecast** refers to the use of time series information in the prediction of the variable of interest. In a time series data set, the information is arranged according to time. For instance, the annual time series of U.S. GDP is the dollar amount of domestic production in United States from 1955 to 2009. For each year, we attach a production dollar amount, and these amounts are reported chronologically starting in 1995 and ending in 2009. Time series data will be the necessary input to construct statistical models and eventually build a time series forecast, which is an estimate of the future value of the variable of interest at a specific date, jointly with a measure of uncertainty. In light of this discussion, it is clear that we can produce an event forecast based on a time series forecast.

[**1.2 Who Are the Users of Forecasts?**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.2)

Planning and preparing for the future requires some forecasting because today’s decisions are functions of what we can foresee today. Broadly speaking, let us think about the economy as a collection of agents—firms, consumers/investors, and government—engaging in productive activities. Why should they be interested in forecasting?

[**1.2.1 Firms**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.2.1)

Although the activities of firms depend on industry specifics, some are common to all of them: capacity investments and capital allocation (size and units), operations planning (personnel, production, inventory, sales, innovation), budgeting (costs and revenues), and marketing (pricing, clients, advertisement). In each of these areas, decisions that depend on the forecasting of the relevant variables must be made. For instance, if a firm is contemplating a business expansion or planning to launch a new product, it must have forecasts of future revenues, new sales, potential new markets, and so forth. Financial firms such as banks and investment companies are in the business of allocating financial capital and managing risk, but to do so, they need forecasts of asset prices (stocks, bonds, exchange rates, swaps, etc.) and of their volatility not only in the domestic market but also in international markets. Every industry has idiosyncratic features and requests forecasts that differ from those of the other industries. You may want to consider different sectors of the economy—energy, pharmaceuticals, technology, financials, information technology, natural commodities, transportation, services—and consider what forecasts would be most relevant to the firms in that industry.

[**1.2.2 Consumers and Investors**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.2.2)

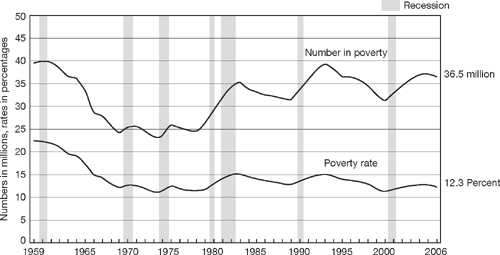
Households play two roles in the economy: They are consumers providing demand for the products and services produced by firms, and they are investors providing savings (in bank deposits, saving accounts, stock and bond purchases, and so on) to the firms to engage in production. In both roles, household decisions are informed by the forecasts of the relevant economic variables. For instance, household revenues depend on the business cycle in the economy, the state of labor markets, and the state of capital markets. For most households, a large proportion of revenues come from labor income, so early decisions about investment in education rely on the prospects of employment. As the life cycle unfolds, capital income becomes important, and household decisions about capital investments rely on forecasts of interest rates, earnings, dividends, and so forth. As consumers and investors, households are concerned about the forecasts of inflation because high inflation will be beneficial if the household is a net borrower or detrimental if the household is a net lender.

[**1.2.3 Government**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.2.3)

Government is a large producer and consumer of forecasts. Many federal and state agencies have armies of professional forecasters concerned about domestic and international macroeconomic variables such as gross domestic product, consumption, investment, exports, imports, employment, prices, interest rates, and exchange rates. These forecasts are key to making major decisions in fiscal and monetary policies. For instance, every month, all economic agents in the U.S. economy wait with anticipation for the statement of the Federal Reserve (Fed) regarding changes in the federal fund rate and guidance on the state of the economy. For all economic agents, this statement is a reading in expectations and is the closest account in modern times to the Oracle of Delphi! This is an excerpt from the Federal Open Market Committee meeting on March 16, 2010:

Although the pace of economic recovery is likely to be moderate for a time, the Committee anticipates a gradual return to higher levels of resource utilization in a context of price stability…. With substantial resource slack continuing to restrain cost pressures and longer-term inflation expectations stable, inflation is likely to be subdued for some time….The Committee will continue to monitor the economic outlook and financial developments and will employ its policy tools as necessary to promote economic recovery and price stability…. The Committee will continue to monitor the economic outlook and financial developments and will employ its policy tools as necessary to promote economic recovery and price stability. (*Source:* htt­p:/­/ww­w.f­ede­ral­res­erv­e.g­ov/­new­sev­ent­s/p­res­s/m­one­tar­y/2­010­031­6a.­htm)

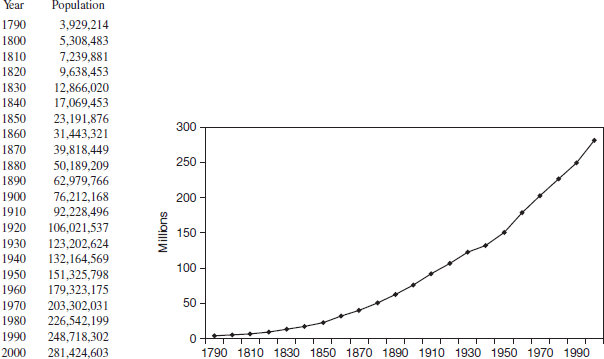
Embedded in this statement are several forecasts: a prediction of resource utilization, which is expected to grow; a prediction of price inflation, which is expected to be moderate; and a prediction of GDP growth, which is expected to be slow. Of course, the timing is an open question, but very likely, the professional forecasters at the Fed produced several forecasts for the short and long terms.



[**1.3 Becoming Familiar with Economic Time Series:  
Features of a Time Series**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.3)

Let us introduce the basic object of analysis in this book: *a time series*. A time series is a sequence of numerical values *ordered* according to time. See the table included in [Figure 1.1](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#fig1_1), which is a time series of the U.S. population from 1790 to 2000 at 10-year intervals.

For each year from 1790 to 2000, we read the population of the United States. For instance, in 1990, the number of people was 248.7 million. A time series plot is a graphical representation of the time series. It always has the same structure: in the horizontal axis, we plot time, and in the vertical axis, we plot the value of the variable. A time series data point is a pair *(t, yt)*. For instance, for *t* = 1990, the U.S. population was about 248.7 million, and we write *y*1990 = 248.7. It is important to visualize time series, particularly business and economic time series, because these are key elements in this book. Visual skills help to foster intuition about the data and to spot features that will guide us in the search for a model. Do not underestimate the value of a graph!



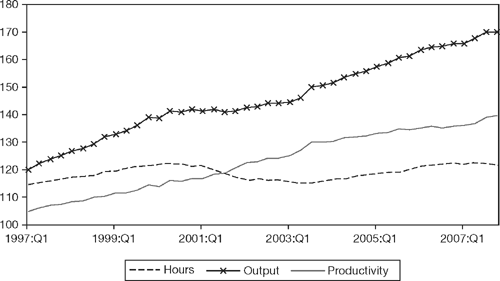
**FIGURE 1.1** Population of the United States at 10-Year Intervals (1790–2000)

*Source:* htt­p:/­/ww­w.c­ens­us.­gov­/co­mpe­ndi­a/s­tat­ab/­cat­s/p­opu­lat­ion­.ht­ml

Some general features in economic time series can be classified within three broad categories: **trends, cycles, and seasonality**. Time series can exhibit one or several of these features. For instance, the time series of U.S. GDP (htt­p:/­/re­sea­rch­.st­lou­isf­ed.­org­/fr­ed2­/) exhibits a general upward trend, but there are also deviations around this trend. Sometimes the time series goes above the trend (indicating expansions) and sometimes below (suggesting recessions) forming a cycle around the trend.

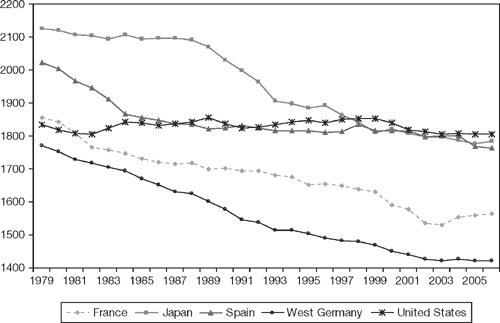
[**1.3.1 Trends**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.3.1)

When a time series evolves slowly and smoothly over time, we say that it shows a *trend,*which is a long-run feature of the data. We should have a great deal of data before identifying any trend. The graph in [Figure 1.1](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#fig1_1) showed a clear upward trend in the U.S. population for the last two centuries. In [Figure 1.2](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#fig1_2), the graph shows an upward trend in U.S. productivity—defined as output per hour worked—and in the output time series. The number of hours worked remained fairly stable for the 10-year period 1997:Q1 to 2007:Q1. As we visualize the time series plots, it is necessary to understand the measurement units of the variables of interest. In [Figure 1.2](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#fig1_2), all variables are reported as indexes. Dealing with an index, we always need the base year to interpret the plot. The base year is the year for which the value of the index is equal to 100. The base year for the series in [Figure 1.2](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#fig1_2) is the first quarter of 1992. Now let us choose any number in the plot, say the output index is 120 in 1997:Q1. Given that the base is in 1992:Q1, the index says that there has been a 20% increase in output from 1992 to 1997 (Q1).



**FIGURE 1.2**  
Quarterly Productivity, Output, and Hours (Non-Farm Business Sector)

*Source:* Bureau of Labor Statistics. Productivity is output per hour. Index, 1st quarter 1992=100.

****

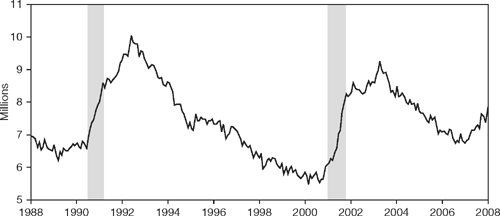
**FIGURE 1.3**  
Annual Hours Worked in the OECD Countries, 1979–2006

*Source:* OECD (2016), Hours worked (indicator). https://data.oecd.org/emp/hours-worked.htm (Accessed on 24 February 2016).

See [Figure 1.3](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#fig1_3) for an example of a downward trend describing the annual hours worked in some Organization for Economic Cooperation and Development (OECD) countries. There is a downward trend in the number of annual hours worked in Japan, Germany, France, and to a lesser extent in Spain. In contrast, hours worked in the United States remained quite stable from 1979–2006. The unit for all variables in [Figure 1.3](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#fig1_3) is the number of hours per year.

[**1.3.2 Cycles**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.3.2)

When a time series exhibits periodic fluctuations, we say that it has a *cycle*. The cycle may be seasonal or nonseasonal. In [Figure 1.4](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#fig1_4), the time series for unemployed persons 1988–2008 exhibits a nonseasonal cycle. Observe that unemployment rises during recessions and drops during expansions creating a cycle, so that every X number of years, we observe similar peaks and troughs. The units of unemployment are millions of unemployed people in a month. [Figure 1.5](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#fig1_5) graphs a nonseasonal cycle in the time series for the number of people in poverty and the poverty rate. These two measures tend to rise in recessions and fall or stabilize during expansions. Both series are pictured within the same graph but have different units. The time series for number in poverty is measured in millions of people per year, and the time series for the poverty rate is measured in percentage as the ratio of the number of people in poverty to the number of people in the United States.

****

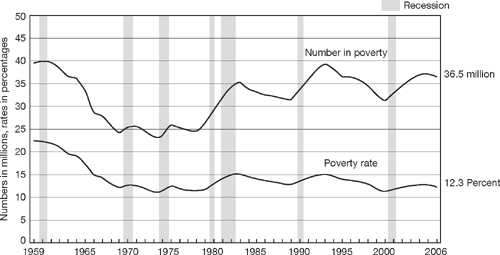
**FIGURE 1.4** Unemployed Persons (Seasonally Adjusted), Monthly Data 1988–2008

*Note:* Shaded areas represent recessions.

*Source:* St. Louis Federal Reserve Bank. FRED.

[**1.3.3 Seasonality**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.3.3)

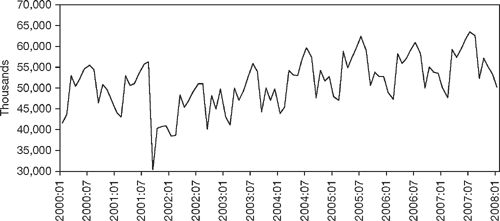
A cycle is *seasonal* when specific fluctuations occur within the calendar year, for instance activities that peak in summer months (or in specific quarters, days, hours, etc.). The time series for revenue passenger enplanements (the total number of passengers boarding an aircraft) in [Figure 1.6](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#fig1_6) exhibits a strong seasonal cycle with pronounced peaks in the summer months of July and August and dips in the winter months of January and February. The measurement unit is 1,000 passengers traveling by plane in a month.



**FIGURE 1.5** Number of People in Poverty and Poverty Rate, Yearly Data 1959–2006

*Note:* The date points are placed at the midpoints of the respective years.

*Source:* U.S. Census Bureau, Current Population Survey, 1960 to 2007 Annual Social and Economic Supplements.

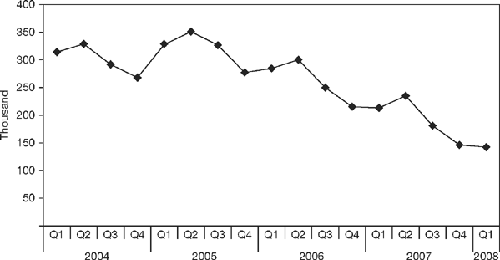


**FIGURE 1.6** Revenue Passenger Emplanements, Monthly Data 2000–2008

*Note:* Total number of passengers boarding an aircraft.

*Source:* Bureau of Transportation Statistics.

Three features, trend, cycle, and seasonal cycle, may come together in a time series. For instance, in [Figure 1.7](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#fig1_7), the time series for new home sales exhibits a seasonal cycle and a trend. Home sales rise during the spring months (Q2) and fall during the winter months (Q4). The seasonal cycle is intertwined with an unambiguous downward trend. In 2004, sales per quarter were about 300,000 houses and in 2008 about 150,000.



**FIGURE 1.7** New Home Sales in the United States, Quarterly Data 2004–2008

*Source:* National Association of Realtors.

For pedagogical reasons, we analyze cycles (seasonal and nonseasonal) and trends separately, but eventually the forecaster’s objective is to produce a forecast for a given time series for which cycles and trends are embedded within each other.

[**1.4 Basic Notation and the Objective of the Forecaster**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.4)

To start talking the language of a professional forecaster, we need to introduce some preliminary jargon and some notation.

[**1.4.1 Basic Notation**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.4.1)

The following table summarizes the forecaster’s basic jargon:

|  |  |  |
| --- | --- | --- |
| **Description** | **Technical name** | **Notation** |
| Object to analyze: | Time series | {*yt*} |
| Value at present time *t:* | Known value of the series | *yt* |
| Future at time *t+h:* | Random variable | *Yt+h* |
| Value at future time *t+h:* | Unknown value of the random variable | *yt+h* |
| Collection of information: | Univariate information set | *It = {y1, y2,…, yt}* |
|  | Multivariate information set | *It = {y1, y2,…, yt, x1, x2,…, xt}* |
| Final objective: | Forecast |  |
|  | 1-step ahead | *ft,1* |
|  | *h*-step ahead | *ft,h* |
| Uncertainty: | Forecast error | *et,h=yt+h−ft,h* |

The data to analyze are a *time series* that is a collection of realizations ordered according to time {*yt*} = {*y*1, *y*2… *yt*…}. When we analyze just one time series, we say that the analysis is *univariate*, and when we analyze several time series *jointly*, for instance two time series {*yt*} = {*y*1, *y*2… *yt*…} and {*xt*} = {*x*1, *x*2… *xt*…}, the analysis is called *multivariate*.

We denote the *present time* by *t*, and *yt* is the known value of the time series at time *t*. The *future* time is denoted by *t* + *h* with *h* being the **forecast horizon** or the step. The future is completely unknown and is represented by a random variable *Yt+h* with many possible realizations or values. As with any random variable, we characterize *Yt+h* by specifying its (conditional) probability density function. A future *unknown* value of this random variable is denoted by *yt+h*. The step *h* could be one minute, day, month, or so on, depending on the frequency of the time series. For instance, if the time series is monthly interest rates and *h* = 2, we are interested in forecasting the 2-month-ahead interest rates. The *h*-step-ahead forecast *ft,h* is based on an **information set**, denoted by *It*, meaning the history or information known to us to the present time *t*. The information set can be *univariate* if it contains only the history of one time series *It* = {*y*0, *y*1, *y*2,… *yt*} or *multivariate* if it contains the information of several time series *It* = {*y*0, *y*1,… *yt*; *x*0, *x*1,… *xt*;…}.

When time passes and we are able to observe the realized value *yt+h*, we can measure how accurate the forecast was. The *h-*step **forecast error** is the difference between the actual realized value and the forecast, that is, *et,h* = *yt+h* − *ft,h*

### [1.4.2 The Forecaster’s Objective](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.4.2)

At the present time t, we wish to forecast Yt+h. The information set is all past history of the series It = {y0, y1, y2, yt}. Based on this information (we say “conditional on the information set”), we will produce three types of forecast:

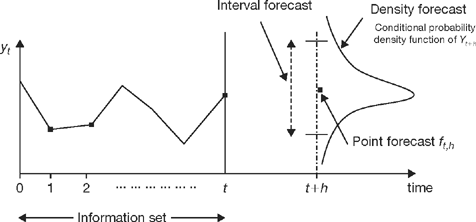
**1.Point forecast** ft,h. This is just a single value. For instance, we may say the 2-month-ahead forecast (h = 2) for the short-term interest rate will be 3%.

**2.Interval forecast.** This is a range of values. We could construct intervals such as (ft,h − kσ, ft,h + kσ) where σ is the (conditional) standard deviation of the random variable Yt+h, and k is a constant related to the probability (or confidence) attached to the interval, which in turn depends on the probability density function of Yt+h. For instance, we may say that in two months, the short-term interest rate will be between 2% and 4% with a 70% confidence.

**3.Density forecast.** This is a probability density function. We could construct the (conditional) probability density function of Yt+h, and by doing that, we will know in a probabilistic sense all future realizations of interest rates. For instance, we may say that in 2 months, the probability for the short-term interest rate to be below 5% is 0.85; that is, P(Yt+h ≤ 5%) = 85%.

See [Figure 1.8](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#fig1_8) for the graphic description of the univariate forecasting problem.

In [Figure 1.8](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#fig1_8), the known information runs up to time t. Based on the information set, we would like to forecast Yt+h at time t + h. At this future time, we picture the (conditional) probability density function of Yt+h to acknowledge that many values of this random variable are possible; some will be more likely than others. The forecast ft,h, constructed at the present time t, is one of these values. In [Figure 1.8](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#fig1_8), we chose the central value of the random variable as a point forecast, and the interval centered on the point forecast as the interval forecast. The final objective in this textbook is to construct three types of forecasts—point, interval, and density—based on the analysis of past information.



**FIGURE 1.8**  
The Forecasting Problem

The multivariate forecasting problem is more complex but conceptually is defined in similar terms to those in the univariate forecasting. Now the time series data set contains more than one time series, say {yt},{xt},{zt}. For instance, we may be interested in forecasting jointly interest rates, output, and money demand. Economic models postulate that output depends on money demand and interest rates; thus, a forecast of output benefits greatly from the information on interest rates and money demand. The information set contains the histories of all time series considered, It = {y1… yt, x1… xt, z1… zt}, and the final objective is to produce the h-step-ahead forecast for the three series corresponding to the future values yt+h, xt+h, zt+h, respectively.

[**1.5 A Road Map for This Forecasting Book**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.5)

This book is concerned with model-based forecasting. We call our models *time series models* meaning that, by analyzing the statistical properties of time series data sets, we can discover the time dependence in the data, and we will express it by means of a mathematical model. If we understand *dependence*, that is, how the present relates to the past, we will be able to project our information from the present to the future.

We proceed to describe broadly the organization of this book. A distinctive feature is that after the introduction of a new concept and/or a new model, an empirical illustration follows. Practical implementation with real data and with simulated data is paramount to the understanding of the most theoretical ideas. We strongly recommend replicating every empirical application because it will help to master these forecasting methods.

The book contains 16 chapters grouped in three large modules. [Module I](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#pt1), [Chapters 1](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#ch1) to [3](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/015_9781315510415_chapter3.html), introduces the field of forecasting by providing context and setting the forecaster’s final objective. Regression analysis is a prerequisite for model-based forecasting, so we offer a brief review of regression techniques. [Chapter 3](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/015_9781315510415_chapter3.html) presents foundational time series concepts, paying special attention to the concept of stationarity. It also provides a fundamental tool of analysis, the autocorrelation functions.

[Module II](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/016_9781315510415_partmoduleII.html#pt2), [Chapters 4](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/016_9781315510415_partmoduleII.html) to [12](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/024_9781315510415_chapter12.html), focuses on forecasting with *linear* time series models introduced across two dimensions: univariate versus multivariate and stationary versus nonstationary. Traditionally, the teaching of forecasting starts with the introduction of trends followed up by the introduction of cycles, which are understood as whatever is left in the series after dealing with the trend. In this book, the pedagogical division is *stationary versus nonstationary* processes. We start by introducing *univariate stationary linear* time series models, which are very well suited to model (nonseasonal and seasonal) cycles in the data. This set of models is characterized as autoregressive and moving average models ([Chapters 6](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/018_9781315510415_chapter6.html) and [7](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/019_9781315510415_chapter7.html)). These models constitute the workhorse of **time series** **forecasting** because they are benchmarks against more complex models should be measured. [Chapter 11](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/023_9781315510415_chapter11.html) introduces *multivariate* versions: vector autoregression. [Chapters 10](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/022_9781315510415_chapter10.html) and [12](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/024_9781315510415_chapter12.html) introduce *nonstationary linear* models, univariate and multivariate, respectively. These models capture trends in the data. We introduce the analysis of trends after the analysis of cycles because we show that a stochastic trend is a limiting case of autoregressive models.

Our final objective is to construct a forecast based on the best model(s) that we can find in the data. To this end, the forecaster faces two important tasks. A priori, before the search for the model starts, the forecaster needs a set of basic tools: a loss function, which drives the optimal forecast; the information set; and the forecast horizon. These are introduced in [Chapter 4](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/016_9781315510415_partmoduleII.html). A posteriori, once the model(s) have been selected, the forecaster needs to evaluate them. In-sample evaluation ([Chapter 8](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/020_9781315510415_chapter8.html)) assesses the logical consistency of the model with the data, and out-of-sample evaluation ([Chapter 9](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/021_9781315510415_chapter9.html)) measures the performance of the model-based forecast. Both [Chapters 8](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/020_9781315510415_chapter8.html) and [9](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/021_9781315510415_chapter9.html) provide forecasting practice based on real time series data.

[Chapter 5](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/017_9781315510415_chapter5.html) offers a link between macro and microeconomic theories and the time series models that we propose. Though understanding this link is not necessary for the construction of a sound forecast, it provides a rich background on how the behavior of consumers, producers, investors, and institutions generates time dependence in the data.

[Module III](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/025_9781315510415_partmoduleIII.html#pt3), [Chapters 13](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/025_9781315510415_partmoduleIII.html) to [16](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/028_9781315510415_chapter16.html), contains advanced material. We introduce more complex dependence across two dimensions. The first extension deals with dynamics of *volatility*. All models in [Module II](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/016_9781315510415_partmoduleII.html#pt2) seek to specify the dynamics of the conditional mean of the process. In [Module III](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/025_9781315510415_partmoduleIII.html#pt3), [Chapters 13](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/025_9781315510415_partmoduleIII.html) and [14](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/026_9781315510415_chapter14.html), we aim to specify the dynamics of the conditional variance. In [Chapter 15](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/027_9781315510415_chapter15.html), we provide several financial applications for which forecasting volatility is crucial. The second extension deals with *nonlinear dependence* in the conditional mean ([Chapter 16](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/028_9781315510415_chapter16.html)). Nonlinear features of time series, such as different dynamics in economic expansions or recessions or different asset dynamics in a bull or in a bear market, require the introduction of more sophisticated models, which are at the frontier of forecasting research.

Between chapters, you will find short sections, “A pause: Where are we and where are we going?” that serve as a compass to navigate the book. We summarize our learning as we go along and set up the next objective. In this way, we position ourselves in the forecasting map so that we do not lose our north.

[**1.6 Resources**](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.6)

*Websites*. Because this textbook offers a hands-on approach, the first step to it is to locate time series data to which to apply the techniques that you will be learning. Fortunately, data are at our fingertips with so many websites offering free access. The following is a list of organizations with plenty of on-line time series information. Our advice is to go each website and become familiar with its content:

htt­p:/­/ww­w.b­loo­mbe­rg.com (daily market data and news)

htt­p:/­/ww­w.b­ea.­gov­/ (U.S. Bureau of Economic Analysis)

htt­p:/­/ww­w.b­ls.­gov­/ (U.S. Bureau of Labor Statistics)

htt­p:/­/ww­w.c­ens­us.­gov/ (U.S. Census Bureau)

htt­p:/­/sd­w.e­cb.­eur­opa.eu (European Central Bank’s data warehouse)

htt­p:/­/ww­w.e­con­oma­gic.com (historical data on economic and financial time series)

htt­p:/­/ww­w.e­rs.­usd­a.gov/ (U.S. Department of Agriculture Economic Research Service)

htt­p:/­/ww­w.f­ede­ral­res­erv­e.g­ov/­rel­eas­es/ (Federal Reserve, Board of Governors Database)

htt­p:/­/fi­nan­ce.­yah­oo.com (current and historical financial data)

htt­p:/­/ww­w.i­mf.­org­/ex­ter­nal­/da­ta.­htm­ (International Monetary Fund—IMF—data)

htt­p:/­/ww­w.f­red­die­mac­.co­m/n­ews­/fi­nan­ce/­ (data and reports on the housing market)

htt­p:/­/mb­a.t­uck­.da­rtm­out­h.e­du/­pag­es/­fac­ult­y/k­en.­fre­nch­/da­ta\_­lib­rar­y.h­tml­ (Kenneth French’s Equity Data Library)

htt­p:/­/ww­w.o­ecd­.or­g/s­tat­spo­rta­l (OECD Statistics Portal)

htt­p:/­/re­sea­rch­.st­lou­isf­ed.­org­/fr­ed2­/ (St. Louis Federal Reserve Bank, Economic Data)

htt­p:/­/uc­laf­ore­cas­t.com (UCLA Forecasting Center)

htt­p:/­/ww­w.w­orl­dba­nk.org/ (World Bank)

htt­p:/­/ep­p.e­uro­sta­t.e­c.e­uro­pa.­eu/­por­tal­/pa­ge/­por­tal­/eu­ros­tat­/ho­me/­ (statistical database of the European Commission)

htt­p:/­/on­lin­e.t­hom­son­reu­ter­s.c­om/­dat­ast­rea­m/ (Thomson Reuters financial statistical database)

htt­p:/­/ww­w.e­ia.­doe­.gov/ (U.S. Energy Information Administration, independent statistics and analysis)

Daily press and specialized magazines. Reading the daily press and specialized magazines is important to grasp how predictions are reported. When you read, take a critical approach to examine the information provided. Ask yourself what information is missing and how you can improve upon the publication. The following is a nonexhaustive list of periodicals:

The Wall Street Journal

Financial Times

The Economist

The New York Times

BusinessWeek (Bloomsberg Businessweek after 2009)

Fortune

Barron’s

Money

Forbes

Academic journals. Forecasting is an active field of academic research. Methodological advances are published first in academic journals. Although the level of exposition in these outlets is highly technical, you may want to browse some issues and read the introduction and the economic and business applications of some articles. You will be surprised how much you will understand once you have studied the material in this textbook. The following is a list of specialized academic journals:

International Journal of Forecasting

Journal of Forecasting

Journal of Time Series Analysis

Journal of Business & Economic Statistics

Journal of Applied Econometrics

Journal of Econometrics

Journal of Financial Econometrics

Review of Economics and Statistics

Software packages. Because the goal of this book is learning by doing, you will need to be proficient in using at least one software package. In the following chapters, you will be exposed to many exercises with real and simulated data that are solved with EViews. This package is very well suited for time series modeling and forecasting, is user friendly, and for most of the chapters, provides “click-and-see” for immediate results. However, the methods in this book can be implemented with any other software. Other great packages for time series analysis are on the market, but they will require some programming skills on your part. The following list contains the most popular software for quantitative analysis:

EViews

GAUSS

Matlab

R

SAS

S-Plus

Stata

## [KEY WORDS](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.7)

Delphi Oracle [p. 1](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#p1)

density forecast [p. 12](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#p12)

event forecast [p. 4](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#p4)

forecast error [p. 12](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#p12)

forecast horizon [p. 11](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#p11)

forecasting [p. 3](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#p3)

information set [p. 12](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#p12)

interval forecast [p. 12](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#p12)

point forecast [p. 12](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#p12)

time series forecast [p. 4](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#p4)

trends, cycles, and seasonality [p. 7](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#p7)

## [EXERCISES](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/004_9781315510415_contents.html#ch1.8)

The following exercises provide a nonexhaustive collection of time series describing some aspects of the U.S. economy. The series are classified within three large categories: the real economy that deals with the generation of income and determination of prices; the financial economy that deals with financial prices such as interest rates, stock indexes, exchange rates; and the social economy that deals with social aspects of the U.S. population. For each time series, please answer the following fundamental questions:

•Definition of the time series: What is it measured?

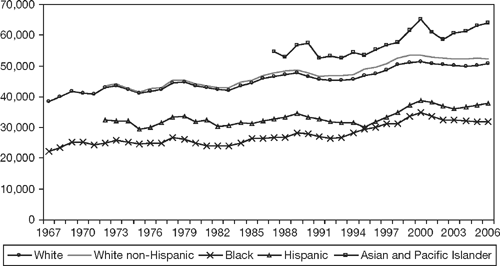
•Measurement units (special attention should be paid to indexes): How is it measured?

•Frequency of the series: What is the periodicity (i.e., daily, monthly)?

•Features of the series (trends, nonseasonal cycle, seasonal cycle): What is the immediate message?

Other specific questions are posed after each time series plot, which will be helpful for comprehension. Micro- and macroeconomic theories as well as U.S. economic history may provide context to these time series.

### Time Series of the U.S. Real Economy



**FIGURE E.1** Median Household Income (2006 Dollars) by Race, Yearly Data 1967– 2006

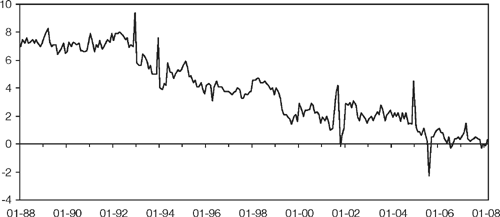
*Source:* U.S. Census Bureau, Current Population Survey.

a. What does “median income” mean?

b. What is the definition of “household” according to the Census Bureau?

c. What is the difference between “real dollars” and “nominal dollars”?

d. What may explain the upward trend for most groups?



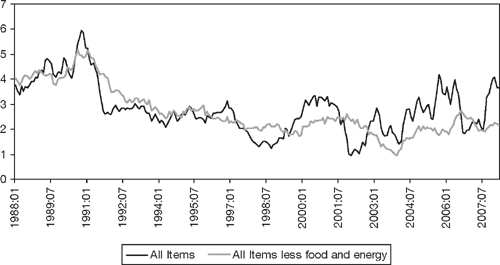
**FIGURE E.2** Saving Rate (%), Monthly Data 1988/01–2008/02

*Source:* St. Louis Federal Reserve Bank. FRED.

a. What is the definition of “saving rate”?

b. What does *negative saving rate* mean, and when does it occur?

c. What is contributing to the downward trend?



**FIGURE E.3** Inflation Rate (12-Month Percentage Change), Monthly Data 1998/01–2008/03

*Source:* St. Louis Federal Reserve Bank. FRED.

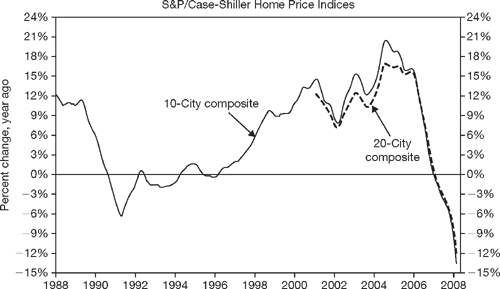
a. What is the definition of *inflation rate*?

b. How would you compute the 12-month percentage change?

c. What is the main difference between the two time series of the inflation rate?

d. What is the purpose of excluding the prices of food and energy?

e. How does the U.S. inflation rate compare to that of other countries?

****

**FIGURE E.4** Home Price Index (12-Month Percentage Change), Monthly Data 1988–2008/02

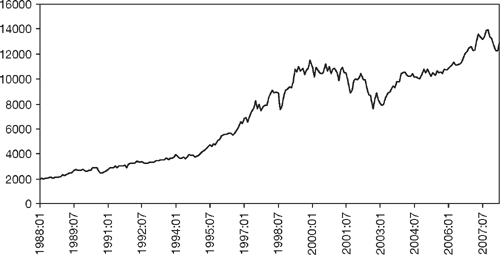
*Source:* Standard and Poor’s and Fiserv.

a. What does the Case-Shiller Index measure?

b. Was there “a housing market bubble” or just an economic cycle from 1996 to 2006?

c. What was going on in the U.S. economy from 1996 to 2006?

**Time Series of the U.S. Financial Economy**

****

**FIGURE E.5** Dow Jones Industrials Index, Monthly Data 1988/01–2008/04

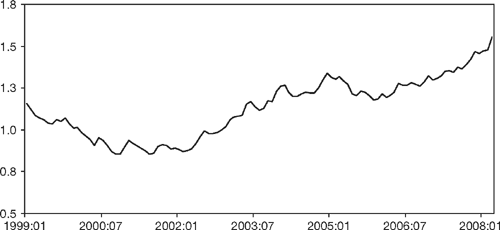
*Source:* Yahoo Finance.

a.What does this index measure?

b. What is the total return over this 20-year period?

c. Why is there a trend?

d. Do similar indexes for other countries exhibit a corresponding trend?



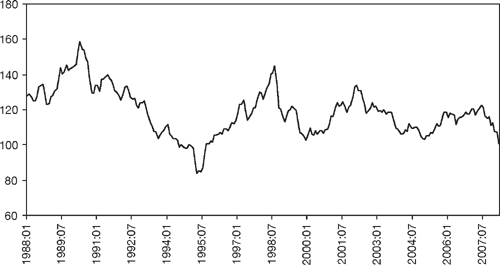
**FIGURE E.6** Exchange Rate U.S.$/Euro, Monthly 1999/01–2008/03

*Source:* St. Louis Federal Reserve Bank. FRED.

a. What is an exchange rate? What does the value of U.S. $/euro = 1.5 mean?

b. In which years is the euro appreciating?

c. In which years is the dollar depreciating?



**FIGURE E.7** Exchange Rate, Japanese Yen/U.S. $, Monthly 1988/01–2008/03

*Source:* St. Louis Federal Reserve Bank. FRED.

a. What does the value of the yen/$ = 120 mean?

b. In which years is the yen appreciating?

c.In which years is the dollar depreciating?

d. Would you be able to construct the series yen/euro with the information in [Figures E.6](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#figE_6) and [E.7](http://e.pub/a77zr9nksotohd4ylbwu.vbk/OEBPS/013_9781315510415_partmoduleI.html#figE_7)?



**FIGURE E.8** Mortgage Rates: 30-Year Fixed Loan, Monthly 1988/01–2008/03

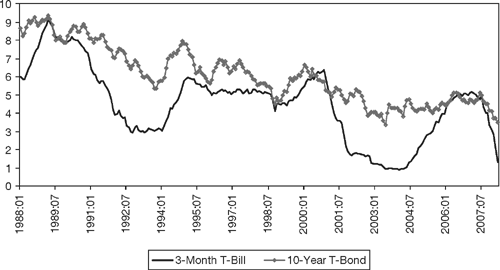
*Source:* St. Louis Federal Reserve Bank. FRED.

a. What is a mortgage rate?

b. What rate should be higher, that of a 30-year or of a 15-year loan?

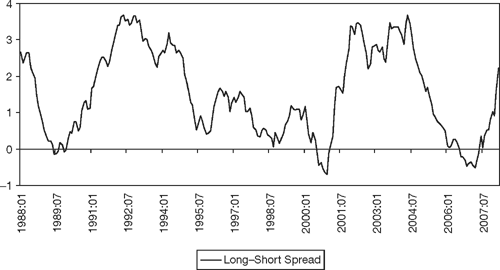
c. What are the economic factors affecting mortgage rates?

d. What is the effect of monetary policy on mortgage rates?



*Source:* St. Louis Federal Reserve Bank. FRED.

**FIGURE E.9** Treasury Rates: Short-Term Rate (3-Month T-Bill) and Long-Term Rate (10-Year T-Bond), Monthly 1988/01–2008/03



**FIGURE E.10** Spread (Long–Short Term Rates), Monthly 1988/01–2008/03

*Source:* St. Louis Federal Reserve Bank. FRED.

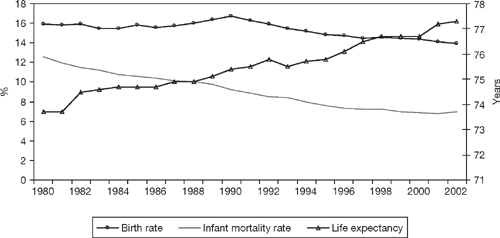
a. What is a Treasury bill?

b. Is the long-term rate always higher than the short-term rate?

c. What is the yield curve?

d. In which years does the spread widen? Why?

**Time Series of the U.S. Society**



**FIGURE E.11** Birth, Infant Mortality, and Life Expectancy, Yearly Data 1980–2002

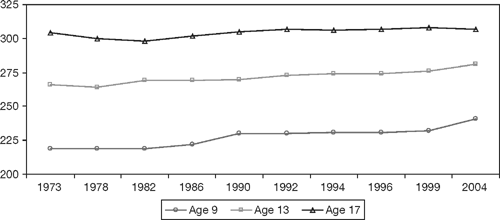
*Source:* National Center for Health Statistics. htt­p:/­/ww­w.c­dc.­gov­/nc­hs

a.What is the meaning of the terms *birth rate* and *mortality rate?*

b. How is *life expectancy* defined?

c. What are the factors that affect all three time series?

d. Why is there an upward trend in life expectancy and a downward trend in infant mortality?



**FIGURE E.12** Mathematics Scores, 1973–2004

*Source:* Department of Education, National Center for Education Statistics. htt­p:/­/nc­es.­ed.­gov­/programs

a. What do these scores mean and measure?

b. Do all three series exhibit an upward trend?