Part I

Institutional Setting

1
Managerial Economics in Public and Nonprofit Administration: An Overview

1.1 INTRODUCTION

Nonprofit and government organizations, justified by custom and tradition, predate the corporate sector historically. They teach, heal, provide social services, and lift people's spirit with recreational, musical, and artistic outlets. These are not flawed businesses to be tolerated whenever the market fails. They have standing in and of themselves, although they cannot show a profit and have difficulty in demonstrating cost-effectiveness, quality, and relevance to long-term Gross National Product.

Any business manager crossing over into the nonprofit and public sector has much to learn. As compared with business executives, nonprofits and public administrators wield less authority, answer to a wider range of shareholders, lack a straightforward performance measure such as profit, are under greater public scrutiny, and have fewer resources to achieve their objectives.

Misunderstandings about the nonprofit and public sectors have serious repercussions. Board members may treat their positions less professionally than their corporate board memberships. Donors neglect to use their influence. Talented administrators, particularly those with business backgrounds, end up frustrated and ineffective because they do not have a handle on the complexity of the nonprofit and public sectors (Silverman and Taliento, 2006). Robert Rubin, former U.S. Treasury Secretary, experienced these differences first hand as described in Application 1.1.

APPLICATION 1.1
Crossing into Government Administration from the Profit-seeking Sector

Robert E. Rubin, U.S. Treasury Secretary under President Clinton, offers several observations about the differences he experienced on becoming a government administrator after 26 years at Goldman Sachs, an investment bank, where he rose to co-senior partner.

Rubin observes that anyone going from head of a large firm in the profit-seeking sector to a cabinet position can be misled by the similarities. In both cases, the person is administrating a large, hierarchically structured organization. However, if one does not recognize the differences, he or she will not be effective in government for the following reasons:

1. In business the chief focus is on profitability. Government has no simple bottom line but a vast array of interest and priorities, some of which are in conflict.
2. Government decision-making is vastly more complex and the decision-making power of the government administrator is more limited. Almost every major decision must be approved

after an extensive interagency process. This permits decision-making power to be far more centralized in profit-seeking firms than in government.
3 A profit-seeking CEO has the power to hire and fire based on performance, to pay bonuses, and to promote capable people. In government, an employee can be dismissed for gross incompetence, but the cost is seldom worth the effort.
4 In government, structural reorganization requires legislation.
5 The key to a top NFP administrator's success is one's relationship with key committees and politicians, and this takes enormous time and thought. Internal oversight of U.S. profit-seeking boards is less intrusive on a day to day basis.

Because of these differences between profit-seeking and government firms, Rubin concludes:

1 Profit-seeking firms have much to offer government firms in improving efficiency in both processes and operations. However, the inherent complexity of government administration remains.
2 In spite of the prejudice many profit-seeking managers have against government administrators, they have much to learn from the way government works. Those who learn how to manage complicated interagency governmental processes have a crucial set of skills less likely to be as well developed in business.

Source: Robert E. Rubin and Jacob Weisberg, *In An Uncertain World*, New York: Random House, 2003.

Like any firm, public and nonprofit organizations create something of value that individuals and groups of individuals are willing to purchase. Public administrators, managing these institutions, wish to avoid politicization and provide the best service for the least cost. However, some organizations are more effective than others. This difference is in part explained by the microeconomic underpinnings of each organization.

With few exceptions, firms producing a good or service are included in one of the following economic sectors:

1 the profit-seeking sector;
2 the government sector; or
3 the nonprofit sector.

This book concerns the second and third sectors, accounting for more than 20 percent of total U.S. employment. Approximately 13 percent of all workers are

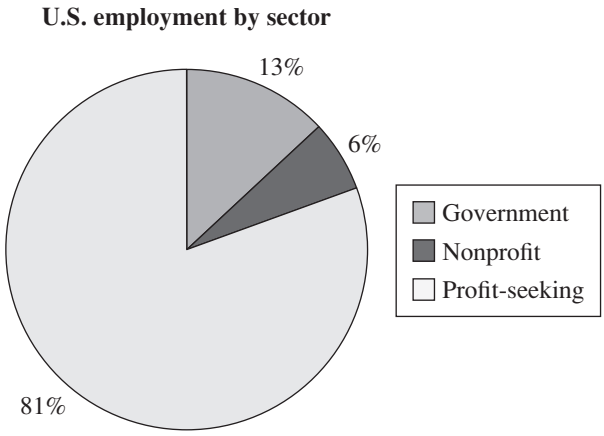

Figure 1.1 U.S. employment by sector (estimated). Precise employment figures by sector are presently unavailable. The sector percentages here are based on Table 1.1 data but adjusted to account for agricultural and self-employed workers excluded from that data. The healthcare industry, for example, spans across the profit-seeking, government, and private nonprofit sectors. U.S. Department of Labor data differentiates between private and public non-agricultural employment but not between the profit-seeking and private nonprofit sectors.

employed by government and 7 percent by nonprofit firms. Figure 1.1 outlines the relative sizes of the three sectors in which the total labor force is employed. Our goal in this book is to show how economics, a social science dealing with scarcity and choice, is relevant to the nonprofit and government sectors.

Consider for a moment the following fantasy. Imagine lounging on a deck chair at your local swimming pool on a sunny summer afternoon. Only the periodic whistle of the lifeguards disturbs the serenity of blue sky and water. Small children in the pool cry "Look at me!" and teens flirt by the concession stand. A few trees and a fence shut out the rest of the world. The only serious decision you face is whether to continue dozing in your chair or plunge into the water for a few laps.

Permit us to draw you back into the world of economics. Most likely the pool of your fantasy is not in the profit-seeking sector, unless located in a private home or hotel. Using valuable real estate for a stand-alone pool is generally inconsistent with profit maximization. A ***profit-seeking firm*** is one that seeks to maximize profit and shareholders wealth. Is the pool of your fantasy operated by the government or by a private nonprofit? Both represent the focus of this text. We use the term ***not-for-profit (NFP)*** to include firms in both the nonprofit and government sectors. A private ***nonprofit firm*** is a non-government organization providing benefits for members or clients within the community. Note that we use the term "firm" for all three types of legal institutions: profit-seeking, nonprofit, and government. In economics, a firm is a producing unit of society, and all three types of institutions create value in the form of goods and services. However, the government and nonprofit sectors share common economic characteristics that differ from firms in the profit-seeking sector in two important ways:

1 NFP firms receive part of their revenue from a source other than from the sale of output to users.
2 These alternate sources of revenue, whether from donations or taxation, affect firm behavior and administrative decisions.

Public administration is the broad study and practice of implementing policy within a government or quasi-government organization. Increasingly, the term public administration implies the management of all NFP firms providing certain goods and services called public goods. This text, designed for government and private nonprofit administrators and students, is less concerned with policy formation but more with policy interpretation and implementation within an organization. Whenever organizations are somewhat autonomous, administrators set fees, decide on output, and make management decisions.

From birth onward, we are in constant association with NFP firms: hospitals, churches, schools, civic and fraternal associations, cooperatives, retirement communities, etc. These institutions are the foundation on which the whole economy rests; they are the necessary socializing institutions that create the environment and legal arrangements under which business transactions take place. The very fabric of society is stitched together with the thread of these firms, public and private. They are not businesses; they neither seek profit nor intend to maximize profits. They are not households. Public and private NFP firms are organized to produce a given outcome, and like any firm they sometimes fail. Many of us were disillusioned with disaster relief provided in 2005 after Hurricane Katrina. It seems as if every sector of the economy was incapable of responding effectively. Some lessons and suggestions for dealing with future catastrophes are discussed in Application 1.2.

APPLICATION 1.2
Lessons from Hurricane Katrina

It is easy to criticize each sector of the U.S., government, private nonprofit, and profit-seeking, associated with disaster relief in the aftermath of the devastation to New Orleans and its outlying areas caused by Hurricane Katrina in 2005. However, Trent Stamp, President of Charity Navigator, suggests that even if the federal government through its Federal Management Agency appears to have failed in its ability to handle the Katrina disaster, the role of private charities in future disasters is more vital than ever.

Lean staffing and volunteer labor limit private nonprofit response. The New Orleans Red Cross office after Katrina was flooded, half its employees lost their homes, and two lost family members. Yet, private charities with donor dollars evacuated residents, served meals, provided cash assistance, and is rebuilding communities. In the midst of the disaster, mistakes were made on the one hand by too hastily

distributing debit cards and providing redundant services and on the other hand by too slowly assisting people in distress.

Stamp suggests that nonprofit and government organizations can learn from their mistakes. He suggests that particular agencies recognize their comparative advantage, refer clients to other charities, cooperate with other agencies rather than be territorial, and refrain from soliciting funds for services for which they have no experience.

Disaster relief procedures should be based on general principles for dealing with unexpected circumstances. After Katrina hit, food, stored 70 miles away in Baton Rouge, failed to reach New Orleans on flooded roads. NFPs must construct worst-case not merely most-likely scenarios. After the September 11, 2001, terrorist attacks, in order to avoid excessive hoarding of funds, organizations initiated "pay as your go" policies in which funds have to be raised and allocated separately for each catastrophe. The unintended consequence was to prevent charities from shifting funds to assist after Hurricane Rita followed Katrina in just four weeks. Thus, it would be imprudent to devise new rules based exclusively on the Katrina experience.

Source: Trent Stamp, "Charities Must Heed the Lessons From Hurricane Katrina," *Chronicle of Philanthropy*, 18 (21), August 17, 2006, 28.

The following three assumptions are the basis of the intersection between the study of economics and NFP firm management:

1 Unlike profit-seeking firms, NFP firms do not focus on what is earned, but rather on what they do. The output of the organization is important independently of income earned by the sale of its output. In fact, the output is important whether or not it earns any income at all.
2 NFP firms operate with internal and external constraints that distinguish them from profit-seeking firms. Although human behavior is a constant, administrators and clients of NFP firms are motivated and rewarded differently than executives and customers of profit-seeking firms.
3 NFP firms allocate scarce resources and create value. As such, they should be operated so as to provide the most benefits at minimum cost. The study of economics provides administrators with tools to do this.

Students earning a Masters in Business Administrations (MBA) expect to apply their training over a range of profit-seeking industries. Of course, some training must be specific to that industry; managing a factory producing steel is different from managing a firm in the fashion industry. Similarly, the experience and training of healthcare administration is different from that of educational administrators, and both are different from social services and the arts. The study of public administration economics respects the ideals and mores of diverse professionals,

their unique rules and regulations, and their standard procedures. Nevertheless, economic theory of a general NFP firm offers insight into how organizations operating across a wide variety of industries can be most effective.

What specifically can economics offer those who study and work in public and nonprofit administration? Public administrators need to internalize five essential economic concepts:

1 The nature of public versus private choice and a clearer understanding of differences and similarities between profit-seeking and NFP firms.
2 A heightened awareness of scarcity with respect to unlimited wants.
3 The primary focus of the organization, with which he or she is associated, is the provision of a specific good or service intended by the sponsor and of value to the client.
4 Individual donors, sponsors, administrators, employees, and clients operate in their own best interests which may or may not be coincident with the organization's goals. Inappropriate incentives produce unintended consequences. Public good provision is constrained by constitutional considerations and "rules of the game" rather than strict economic efficiency.
5 Each managerial decision has a cost measured by the value of the foregone alternative; every choice, then, can either move an organization toward or away from furthering its purpose.

The best way to master these core economic concepts is through direct experience in the management of various NFP firms. But that would take a thousand lifetimes. By explaining and applying economic concepts to a variety of nonprofit situations, this text aspires to assist those in government and private NFP firms in the complex task of effectively providing social goods and services. The economic tools presented in this text do not require advanced mathematics, such as calculus. Nevertheless, understanding simple graphs and equations is essential, as is a calculator for solving end-of-chapter exercises. The Appendix to this chapter is a tutorial on regression, a statistical tool employed in many of the Application presented in each chapter. Any statistical software program, such as Excel, is helpful but not required to understand the tutorial.

In this text, NFP issues and applications are fully incorporated with economic theory: indeed, the text follows the traditional microeconomic format as shown in the following part numbers:

I Institutional Setting
II Consumer Theory and Public Goods
III Production Theory and Public Administration
IV Input Markets and Cost–Benefit Analysis.

Gross Domestic Product (GDP) is the market value of all final goods and services produced within a country in the course of one year in all three sectors: government, private nonprofit, and profit-seeking. What share of GDP is produced in the two NFP sectors, the focus of this book? At present, we are unable to answer that

Table 1.1 Annual average total of persons, 16 years and over, employed as wage and salary workers in the United States, 2002 (thousands)

Industries	*Total workers*	*Private sector*	*Government sector*	*% of total*	*Private sector*	*Government sector*
Total employment, 16 years and over	126,438	106,750	19,689	100.0	100.0	100.0
Agriculture	1,282	1,229	53	1.0	1.2	0.3
Mining	490	488	1	0.4	0.5	0.0
Construction	8,367	7,899	469	6.6	7.4	2.4
Manufacturing	16,918	16,848	59	13.4	15.8	0.3
Wholesale and retail trade	18,625	18,536	89	14.7	17.4	0.5
Transportation and utilities	6,872	5,318	1,554	5.4	5.0	7.9
Information	3,545	3,386	159	2.8	3.2	0.8
Financial activities	8,881	8,699	182	7.0	8.1	0.9
Professional and business services	12,138	11,816	322	9.6	11.1	1.6
Education and health services	26,493	16,429	10,064	21.0	15.4	51.1
Educational services	11,541	3,125	8,415	9.1	2.9	42.7
Hospitals	5,321	4,661	660	4.2	4.4	3.4
Healthcare, non-hospital	7,357	6,830	527	5.8	6.4	2.7
Social assistance	2,274	1,812	462	1.8	1.7	2.3
Leisure and hospitality	10,907	10,519	388	8.6	9.9	2.0
Other services	5,613	4,827	30	4.4	4.5	0.2
Repair and maintenance	1,518	1,509	9	1.2	1.4	0.0
Personal and laundry	1,475	1,453	21	1.2	1.4	0.1
Membership associations and organizations	1,864	1,864	0	1.5	1.7	0.0
Private household	757	757	0	0.6	0.7	0.0
Public administration	6,307		6,307	5.0	0.0	32.0

Self-employed and unpaid family workers are excluded. Private nonprofit employment with respect to total nonagricultual private firm employment is 7.8 percent. A breakdown by industry for private nonprofit and for-profit employment is not available.
Source: Based on current population survey.

question precisely. The amount of government spending, not used to purchase goods and services, is generally available and is an approximate measure of the value of output produced by government employees. However, no similar figure exists for nonprofits in the United States.

Labor-force data, for those presently holding jobs or seeking work, is usually broken down by industry, but not by sector. We know, for example, the number of nurses working in various industries but not if the firm in which they are employed

is a nonprofit or profit-seeking corporation. Therefore, we can merely estimate changes and the amount of workers in the NFP sectors. The Johns Hopkins Center for Civil Society and the United Nations Statistical Division have worked together to address this issue. Together they have issued the *Handbook on Non-profit Institutions in the System of National Accounts*.

The Urban Institute also collects data on the size and characteristics of the private nonprofit sector in the United States. The United States Department of Labor provides data on employment broken down into industries and into private versus government workers. As of now, the government does not survey private nonprofit employment by industries. Based on information provided in Table 1.1, we can make some generalizations about employment in both the private nonprofit and government sectors and the relative sizes of those sectors.

The United States government employs about 16 percent of the civilian labor force, and private nonprofits account for 7.8 percent of nonagricultural private employment. Thus, over 20 percent of all U.S. civilian employees work in either the government or in the private nonprofit sector. Half of all government workers are in education and health services and 32 percent in public administration. Another 4.3 percent of government workers are in social service or leisure/hospitality industries. Education, health services, and religious ministry dominate the private nonprofit sector, followed by social services.

In Europe, it is traditional to speak of the "social sector" including cooperative organizations, such as unions, buyer and seller co-ops, worker owned firms, and credit unions, as part of the private nonprofit sector. Cooperatives play a less significant role in the United States, because the distribution of net income (revenue minus costs) to those associated with tax exempt nonprofits is prohibited. In all economically advanced economies, the private nonprofit sector tends to be large and is increasingly regarded as a type of innovative public management. Therefore, this book focuses on private nonprofit as well as government management.

1.2 ECONOMIC THINKING IN NOT-FOR-PROFIT ADMINISTRATION

Often students resist abstract graphical and mathematical models that stylize, compare, and generalize. What do drug treatment programs have in common with automobile production? Aren't differences between NFP and profit-seeking firms obvious and unique? What benefit does rigorous abstract analysis offer to practicing administrators? The simple answer is that an abstract treatment counters generalizations about NFP firms that are wrong, misleading, and dangerous. Consider just a few of these generalizations:

- All large bureaucracies, profit-seeking or not, operate similarly.
- Certain goods and services, such as education or medicine, are best provided in NFP firms.
- NFP firms are under-financed.
- Medical care and education are basic necessities and human rights
- NFP administrators cannot assess or measure effectiveness

- NFP firms do not experience the highs and lows of the business cycle.
- NFP firms lack "bottom line" discipline and are therefore less efficient that profit-seeking firms.
- Competitive individuals go into business and humanitarians are drawn into the NFP sectors.
- NFP firms should treat clients like customers.
- Clients are always incapable of determining quality.

Some of the above generalizations are clearly wrong; most contain kernels of truth needing qualification. As you examine the nonprofit environment more carefully, you may come to find the following economic assumptions more useful:

- The institutional and legal structure of a firm matters.
- Public goods, such as safety, and quasi-public goods, such as education and medicine, are exchanged in a wide network of firms in all three sectors.
- Each NFP, given a fixed budget, can choose to maximize output and/or provide a better product.
- A profit-seeking firm is intended to maximize shareholder wealth. A nonprofit firm is intended to further the goals of its sponsoring agency.
- Employees, regardless of the type of employer, are similarly motivated responding to rewards and incentives.
- Administrators interpret NFP policy choosing what, how, and for whom to produce.
- NFP clients do not pay "out of pocket" full cost; they are subsidized.

A critical difference between profit-seeking and NFP firms is the ***non-distributional constraint***. If a firm claiming nonprofit tax status generates an amount of revenue exceeding costs, it may not distribute that surplus to individuals who own or control the organization. Clients are not entitled to rebates and administrators do not get bonuses based on financial performance. Budget surpluses in government are certainly not divided among employees and in most instances are forfeited to make up deficits elsewhere.

The term NFP refers to privately sponsored nonprofit firms as well as government agencies. We exclude all profit-seeking firms from our analysis including those operating in the education, healthcare, and social service industries. We exclude as well, all profit-seeking subsidiaries of nonprofit or government organizations. Included as NFP firms are all government units, nonprofit and state colleges and universities, healthcare organizations, voluntary health and welfare organizations, foundations, and other nonprofit professional, trade, scientific, and religious organizations. Whenever necessary, the term "private nonprofits" is used to distinguish these firms from government agencies.

1.3 ECONOMIC SCARCITY

Economics is the study of scarcity and choice because resources are limited. Every nonprofit firm, public or private, uses available scarce resources to create value.

Economics ***resources*** are the natural, physical, and human inputs used to produce goods and services; sometimes resources are referred to as inputs or factors of production. The amount of resources available for production is finite. Imagine if you will that the residents of a country decide that each person had a right to dental care. The goal is to provide dental care such that each permanent tooth of every resident is intact at death. All known cosmetic and orthodontic services are included. Estimate in thousands of dollars what this service would cost per household at current prices and compare this with the pre-tax salary of the average worker. However worthy the goal, the most affluent countries in the world could not provide this service without foregoing significant other services.

Economics and value

Economists are criticized as knowing the price of everything and the value of nothing. This is humorous but untrue. Value is at the heart of economics. The ***opportunity cost*** of anything equals the highest value of what one gives up in attaining that product. The opportunity costs of a 3-credit university class is, for example, $2,000 in tuition, $150 for textbooks, and $3,000 of foregone employment income missed through class and study time. The opportunity cost is $5,150 or $1,712 per credit hour. The expected stream of income or satisfaction is the best measure of benefits. The opportunity cost of a degree in public policy, for example, equals "out of pocket" tuition, books, and other fees *plus* any income given up in the process of attaining it. The benefits are the expected increases in income and/or satisfaction earned as a result of having the degree. Admittedly, individuals, firms, and societies often do not have clear information in assessing the net benefits (benefits minus costs) of what something is worth. You, at best, only have a general idea of the value of your time in taking this class and its future expected benefits.

Assume, however, that one does have a fairly accurate estimation of full costs and benefits. In certain situations, one does not have an incentive to reveal the net value (costs minus benefits) to anyone else. Consider, for example, that you know the extent to which a degree will increase your life-time earnings. You may wish to withhold this information. The financial aid officer, if he or she knew the value of your future earnings, is less likely to offer scholarship assistance, saving the taxpayer or donor the expense of subsidizing your education. Individuals at times have an incentive to be less than candid about costs and benefits; therefore, economists observe not what people say but what they do.

Economists study the market, a stage on which buyers and seller meet, to provide information about the value that buyers and sellers place on a product. NFP administrators need to estimate client and sponsor willingness to spend for each level of service provided. In addition, he or she needs to calculate what the firm is capable of providing at difference levels of support. Subsidized clients desire the highest level of educational and medical services available and, if they are able, subsidize quality at personal expense. Parents, for example, supplement school-provided music training with privately paid lessons. Administrators may wish to provide a high level of care, given best practices. Due to budget constraints, they compromise. Scarcity affects consumer choice studied in Part II and producer choice as presented in Part III.

Market prices do not always accurately reflect society's costs and benefits. For example, the costs to society of environmental pollution are underestimated, as are the benefits of childhood inoculations. Laws and regulations are needed to correct for this type of failure; policy development is not treated in detail in this text. However, when markets fail and profit-seeking firms fail to meet people's needs, NFP firms, public and private, come to the rescue. Decision-making within NFP firms is the focus of this text. Market failure is shown to be a primary justification for nonprofit and government provision in Chapter 5.

Ordinary markets fail to measure the value of public and quasi-public goods, such as security, safety, and park conservation. It is through voting, politics, and cooperative association that we express our desire for these types of goods. Chapter 3 outlines in detail the basic economic distinction between private and public goods. Because NFP firms specialize in providing these services, Chapter 5 analyzes collective provision of public and quasi-public goods.

1.4 ECONOMIC INSTITUTIONS

Macroeconomics (the forest) is concerned with the aggregate economic well-being of a particular region or country. It addresses such topics as GDP, unemployment, inflation, poverty, economic growth, government spending, and taxation. Macroeconomics assists decision-makers in formulating policy. Microeconomics (the trees) is concerned with individual firms rather than aggregates of firms, individual households rather than total populations, and specific industries rather than total output. Microeconomics assists decision-makers within economic institutions. This text is essentially microeconomics for those studying or working with NFP firms. The assumption is that public administrators do not make policy but rather implement it. Another assumption is that nonprofit administrators do not attempt to change society, but provide specific services which benefit recipients.

Consider the macro versus micro orientation of a nonprofit interest group, such as the American Association of Retired Persons (AARP) or the National Association for the Advancement of Colored People (NAACP), or the Sierra Club, the oldest American environmental organization, founded by John Muir. The expressed goal of these organizations is to affect national policy and hence they are macroeconomic in orientation. What we emphasize in this text, however, is that all administrators deal with the following microeconomic questions:

1. What combinations and how much of each service should they produce?
2. How should they allocate tax revenue, donations, fees, and other sources of income?
3. What technique should be used in producing these services?
4. How can they minimize costs for any given level of output?
5. To whom should benefits of this output be directed?

Economics, as a discipline, has developed a general theory of the firm, relevant to administrators wrestling with the above questions. Economics provides models and a standard vocabulary for discussing these issues. This framework can be abstract

and generic, but, if correctly applied, it highlights the uniqueness and complexity of a particular task. The intention is to provide theoretical and practical assistance to those responsible for a firm's effectiveness.

Following Chapters 2's analysis of the role of the NFP sectors as a whole, subsequent chapters somewhat mirror or resemble a standard microeconomic managerial course. Supply and demand are introduced in Chapter 3. The optimization techniques, presented in Chapters 4 and 7, are microeconomic concepts. A framework for analyzing costs for a given output is provided in Chapter 6, and the optimal combination of labor and capital is addressed in Part IV.

1.5 ECONOMIC BEHAVIOR

Rational self-interest is the basic simplifying economic assumption of human behavior; individuals operate rationally in their own interest. This may seem overly cynical, especially with dealing with NFP firms. Yet, it is quite the opposite. Individuals are not assumed to be predatory; just aware of their alternatives. Therefore, they tend to make clear and consistent choices. Rational self-interest does *not* imply either of the following: (1) one's happiness increases as another's decreases; (2) individuals are fundamentally materialistic. Economists merely assume that individuals, including donors, sponsors, employees, and clients of all firms, define their particular preferences, including philanthropy, and act on them.

Self-interest does not preclude cooperation. Indeed, without cooperation government and nonprofit firms could not exist. Application 1.3 makes the case for ***conditional cooperators***, individuals willing to initiate cooperative behavior anticipating that others will reciprocate.

APPLICATION 1.3
Fool Me Once, Shame on You; Fool Me Twice, Shame on Me!

Economists start with the following premise about human behavior. People are rational; they operate in their own self-interest and generally make clear and consistent choices. But suppose everyone were a pure forward-looking *rational egotist* making decisions based solely on individual benefits. Would NFP firms exist? Why should I donate or tolerate tax increases, if the benefits do not accrue to me personally? Indeed there is some support that certain rational individuals are not likely to cooperate in certain settings, even when such cooperation would be to their mutual benefit. What then are the circumstances that might induce one to contribute given the uncertainty of receiving anything in return?

After surveying the literature on cooperation in various experiments, Elinor Ostrom of Indiana University argues that along with

self-interest some individuals bring with them a set of norms and values that support cooperation. In one study of a pool of 136 subjects, 40 percent ranked cooperative outcomes higher than other outcomes even when their individual payoff was lower. In games where one player divides a sum of money and the other player must accept the division, "nice players," who give away at least 30 percent of the funds, tend to choose cooperative strategies in subsequent games. Economists note, however, that cooperative behavior tends to decline in situations when individuals realize that they could win more by defecting. In subsequent rounds of experimental games in which players experience multiple instances of betrayal, cooperation and reciprocity diminish but are not completely eliminated.

Nonprofit firms are dependent on donations, and government agencies rely on general tax revenue. Therefore, certain individuals at times are required to or freely volunteer to behave in ways other than in rational self-interest. Ostrom makes the cases for *conditional cooperators* as individuals willing to initiate cooperative behavior when they think others will reciprocate. These individuals continue to cooperate as long as a sufficient proportion of others reciprocates. You, for example, might be willing to initiate parties in your college dorm or apartment building and will continue to participate as long as you receive help and support from other residents. Ostrom identifies *willing punishers* who, if given the opportunity, will chastise or fine those who fail to cooperate. Their sense of fair play makes them willing to reward those who have contributed more than the minimal level. For example, *willing punishers* may argue that those who were willing to setup the party should at least be able to choose the music!

In any given society, *conditional cooperators* and *willing punishers* permit *rational egotists* to form institutions, such as nonprofits and government agencies, in order to provide goods and services that unsubsidized markets fail to deliver.

Source: Elinor Ostrom, "Collective Action and the Evolution of Social Norms," *Journal of Economic Perspectives*, summer, 2000, 37–158.

Psychologists and anthropologists have more sophisticated models of human behavior and preference formation, but economic methodology prefers to take individual preferences as given. Therefore, NFP clients, as analyzed in Chapter 3, seek the most service at minimal personal cost. Chapters 8 and 9 show how donors/volunteers and sponsors subsidizing a firm do so with certain goals in mind; they expect to attain the highest outcome for each hour volunteered or dollar donated.

Certain NFP administrators may be less competitive or more humanitarian than profit-seeking colleagues, but this stereotype does little to explain NFP behavior. Most employees work to provide for themselves and their families and to further

their careers. Chapter 7 discusses conflicts between managerial behavior and firm mission. To prevent subversion by individuals of the NFP firm's primary purpose, Chapter 10 introduces a technique, called cost benefit analysis. Cost benefit analysis tests the extent to which proposed projects achieve appropriate outcomes.

1.6 ECONOMIC CHOICE AND OPPORTUNITY COSTS

Scarce natural resources, physical and mental labor, capital consisting of machinery and plant, and managerial know-how are combined to produce goods and services. NFP firms contract for resources in competitive markets in competition with profit-seeking firms. Government and nonprofit firms have to purchase their paper, computers, and computer programming services as all firms do by being willing and able to pay the going market price. The organization then decides how and for what these resources should be used. NFP administrators play an important role in allocating society's scarce resources.

No firm can survive in the long run if its expenses consistently exceed its revenue, and every worthwhile project involves costs that could be used to produce something else. In this text, we are concerned with choices made within NFP firms about what programs to introduce, to continue, to cut back, or to eliminate. The best way to approach this decision is to ask, "What is the highest value that could be obtained if the resources used in this program were shifted into producing something else?" The real or opportunity costs of a program are other programs forfeited, because the firm does not have the resources to do everything. The real cost is what, for example, a university gives up in expanding student slots in one course of study by reducing slots in another.

Suppose the marginal or incremental opportunity cost of training one more dental hygienist is a reduction of two classical scholars. Assume that it takes the university twice as many resources to train a dental hygienist as it does a classical scholar. If it is less costly to produce classical scholars, why would a university train hygienists? Perhaps students, donors, and the state are willing to pay more than twice as much for hygienists as they are for classical scholars. The university, then, could be indifferent between the two programs. Note carefully that we are not saying that the university should be indifferent. The university may wish to subsidize classical studies. What we are suggesting is that resource allocation within a nonprofit firm reflects the desired outcomes of those making decisions. However, we wish to note that, like profit-seeking firms, the competitive external environment, discussed in Chapters 8 and 9 of this book, explains much of what we observe in NFP choice.

Production possibilities and scarcity

Figure 1.2 assumes a fixed annual budget for a university. This budget is allocated between thousands of undergraduate credit hours on the Y axis and thousands of graduate credit hours on the X axis. Along the linear production possibility frontier in Figure 1.2, if graduate credit hours are expanded, then fewer resources are

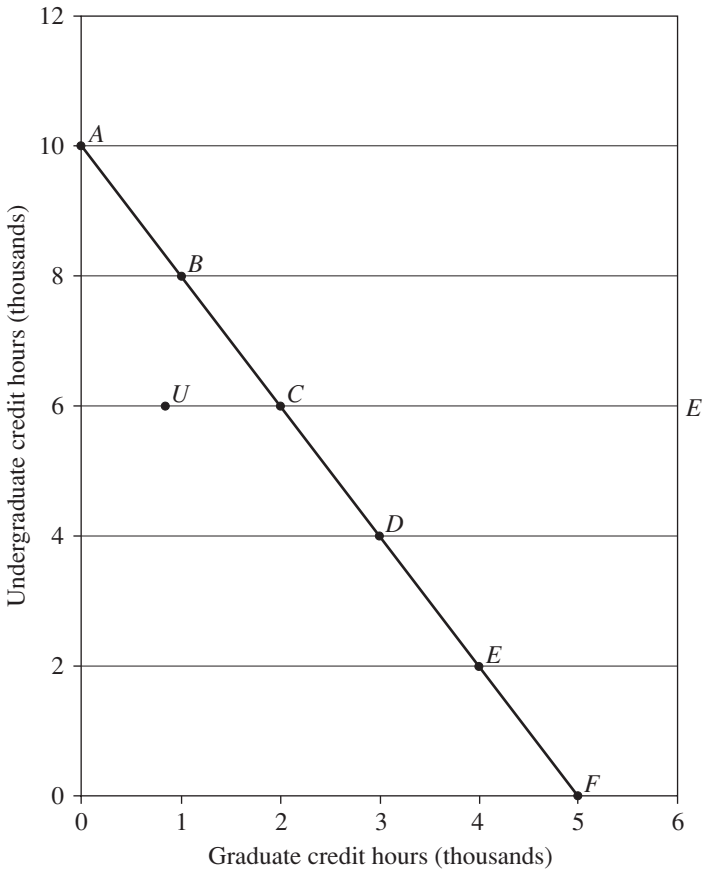

Figure 1.2 Possible combinations with a fixed budget: the production possibilities frontier. With a fixed budget, a university can provide a finite number of graduate and undergraduate credits. As graduate credits increase, undergraduate credits decline. The linear frontier in this figure assumes that the opportunity cost for each incremental increase in graduate credit is two undergraduate credits. For every graduate credit offered, the institution has an opportunity cost or forfeits providing two undergraduate credits.

allocated to undergraduates. Each incremental expansion of graduate credit hours by 1,000 reduces undergraduate credits by 2,000. The opportunity cost of each unit of *X* is two units of *Y*, suggesting that it costs the university twice as much to provide graduate credit as it does undergraduate credit.

In the linear production possibilities frontier presented in Figure 1.2, graduate credit is expanded one by one from 0 units of *X* to 5 units of *X*, until the budget is exhausted and 0 units of undergraduate education are supplied. This frontier demonstrates efficiency and cost. Productive efficiency is represented along the frontier as maximum possible production given present budgetary constraints. Costs are represented by the extent to which output of one program must be decreased to produce one more increment of the other program.

Productive and allocative efficiency

A firm achieves ***productive efficiency*** when maximum product is produced at least cost. Consider, for example, a firm operating at point *U* in Figure 1.2. It is producing 1 units of *X* and 6 units of *Y*, yet it is capable of producing more output with its fixed budget. Each combinations of *X* and *Y* along the frontier represents productive efficiency. All points, *A* through *F*, are equally efficient given existing technology. ***Technology*** is defined as knowledge about how best to produce a given product; an increase in technology reducing the cost of supplying undergraduate credit raises the intersection of the budget line along the *Y* axis. Application 1.3 discusses the controversial issue of the possibility of increasing productivity in service industries found in the NFP sectors. Assuming that the best available technology is in place and represented on the frontier, administers maximize output somewhere along the possibility line for a given budget. An increase in the university's budget or technology pushes the frontier out to the right, and, a budget decrease moves it inward.

APPLICATION 1.4
Can Government and Nonprofit Firms Become More Efficient in Providing Services?

NFP professionals are skeptical of increasing efficiency. They claim that economic models dealing with tangible products like refrigerators and automobiles are irrelevant in providing services, such as education and healthcare. Can a single first-grade teacher be expected to teach more than 25 students per year without decreasing quality? Can a single critical-care nurse tend to more than 6 patients per shift without decreasing quality? Perhaps, not. This does not, however, mean that services for existing clients cannot be improved.

Many profit-seeking firms, such as airlines and financial institutions, provide services. Yet, they have implemented technology to become more efficient and reduce costs. General Electric Co. and healthcare provider Intermountain Health Care have entered into an agreement to develop an internet computer package providing physicians access to clinical best practices at the moment of treatment.

Presently, it is very expensive to fully computerize the healthcare system to include newer best practice technology or even existing technology. Electronic medical records eliminate handwritten orders contributing to errors. Hospitals and physicians are reluctant to invest in technology as the best use of the scarce medical funding.

Do you believe that industries providing services can become more efficient? How would you go about determining if it is worthwhile for a healthcare facility to adopt computerized technology?

Source: Kathryn Kranhold, "GE, Nonprofit Plan Tool for Physicians," *Wall Street Journal*, February 17, 2005, D3.

Where on the production frontier should a firm produce? For example, how should a university decide between graduate and undergraduate programs? The answer varies and is normative, depending on the values of those involved. Some prefer all graduate programs, other all undergraduate, and still others some combination of *X* and *Y*. A firm achieves ***allocative efficiency*** if the budget is incrementally distributed to whatever program yields the highest sum of value for those sponsoring the project. If a nonprofit firm produces at *C* rather than *B*, allocative efficiency requires that the value to the firm's donors and clients of an additional unit of *X* be worth more than the loss of 2 units of *Y*. The real challenge in NFP provision is to estimate value when clients are not paying full cost.

Consider reasons why a private university might moves from points *B* to *C* in Figure 1.2. First, a donor preferring *X* to *Y* may leverage a small or large donation to influence the move. Second, nonprofit administrators or senior faculty, depending on their interpretation of the university's mission, could shift resources to project *X*. Finally, students by their willingness to pay higher tuition for each unit of *X* can influence the change. Tuition is best understood as a fee rather than a price. User fees, paid by clients of NFP firms, generally fail to cover full cost. Similarly, any government-sponsored university, guaranteed a subsidy per credit hour from the state, has an incentive to direct more resources to graduate education if graduate students are willing to pay higher user fees than undergraduates.

Is there any objective way to determine if a move from *B* to *C* is "in the public interest"? In practice, net benefits are subjective and cannot be precisely measured. Economists attempt to frame the question theoretically by applying the ***Pareto Criterion***. Pareto, a nineteenth-century Italian economist, suggested that a move from *B* to *C* is desirable if no one in society is worse off and at least one person is better off. Furthermore, the move should not take place, unless the benefits from an increase in *X* are sufficient to compensate those harmed by decreased *Y*. Because clients do not pay full cost, we cannot prove that present and future benefits of increased graduate credit hours shown in Figure 1.2 exceed the cost of reducing undergraduate hours. However, if we observe such a move, this is indeed what administrators intend to indicate.

There are five immediate advantages of adopting an economic approach to NFP administration.

1. It focuses on the differences between for-profit and NFP firms.
2. Budgetary constraints are made explicit. Choices must be made between alternative goals.
3. Optimization is stressed, meaning the practice of either (a) minimizing costs for a given level of output, or (b) increasing output with the same amount of resources.
4. Costly efforts to change human behavior are replaced with effective incentives.
5. NFP outcomes are viewed as a process whereby individuals filter their preferences through private markets as well as through collective and political choice. Parents, for example, provide schooling for their child by a combination of selecting to live in a certain school district, paying tuition, electing pro-education government officials, and participating in home-school associations.

Present and future administrators, regardless of their academic major and former resistance to economics, need a framework for analyzing various types of NFP firms and the industries in which they operate. Decision-makers in the NFP environment confront "what happens if?" type questions. The "if" may be a change in government tax policy, a shortfall in expected subsidies, or increased competition. Microeconomics tools assist managers in responding to "what happens if" type questions. While the tools of microeconomics are more commonly applied in the profit-seeking sector, the economists' toolkit is useful as well to government and nonprofit administrators.

CONCLUDING NOTES

- Public and private not-for-profit (NFP) firms neither seek to profit nor intend to maximize profits. They are organized to produce a given outcome independently of income earned by the sale of their products.
- NFP firms are the dominant socializing institutions of all societies creating the environment and legal arrangements under which most transactions and all business take place.
- Private nonprofit firms and government agencies have institutional and legal structures affecting their behavior. They often engage in supplying services inadequately provided in profit-seeking markets.
- The non-distributional constraint stipulates that a NFP firm may not distribute any surplus to individuals who own or control the organization.
- Economics is the study of scarcity and choice given limited resources.
- Microeconomics, sometimes called managerial economics, provides nonprofit and government administrators with a framework for decision-making.
- Individuals, including donors, sponsors, employees, and clients of NFP firms define their preferences and act on those preferences.
- All NFP firms are constrained by revenue. Over time, they cannot spend more than clients, sponsors, taxpayers, and donors are willing to pay.
- The real or opportunity cost of a NFP firm producing a particular good or service is the value of the alternative products that could have been produced.
- Productive efficiency requires that society's scarce resources be used to produce at lowest possible cost. Allocative efficiency requires that society's resources be used to produce those goods and services most desired by persons contributing to or paying fees to obtain the product.
- Clients of public and nonprofit agencies generally do not pay full cost for the services received. Client fees do not cover all costs.
- Economists use the Pareto criterion to examine if a move from situation B to situation A is beneficial. The move is considered desirable if no one in society is worse off as a result of the move and at least one person is better off.

KEY TERMS

Profit-seeking firm
Not-for-profit firm (NFP)
Nonprofit firm
Non-distributional constraint
Resources
Opportunity cost
Macroeconomics
Microeconomics
Rational self-interest
Conditional cooperators
Productive efficiency
Technology
Allocative efficiency
Pareto Criterion

SUGGESTED READINGS

McConnell, Campbell R., and Brue, Stanley L. 2002: *Microeconomics: Principles, Problems, and Policies*, New York: McGraw-Hill.

Silverman, Les and Taliento, Lynn 2006: "What Business Execs Don't Know – but Should – about Nonprofits," *Stanford Social Innovation Review*, summer, 37–43.

Young, Dennis R. and Steinberg, Richard 1995: *Economics for Nonprofit Managers*, New York: Foundations Center.

END OF CHAPTER EXERCISES

Exercise 1.1

Explain how the non-distributional constraint affects or does not affect the following organizations: the U.S. Department of Commerce, Hospital Corp. of America, the Teamsters, Sylvan Learning Centers, Indiana State Lottery, the National Football League, Main Street Cardiology Partnership, and your University bookstore.

Exercise 1.2

Discuss how the following individuals, with respect to each organization listed in Exercise 1, operate rationally in their own self-interest: owners/stockholders, donors, administrators/managers, employees, volunteers, customers, and clients. Does rational self-interest preclude cooperation?

Exercise 1.3

The following represents production possibilities of clients served in a hospital with a fixed amount of capacity, staff, and equipment:

Type of production: (in thousands of clients billed in one year)	*A*	*B*	*C*	*D*	*E*
In-patients	0	2	4	6	8
Out-patients	30	27	21	12	0

a. Graph the information presented in the above table. (Hint: the production possibilities curve is curved because opportunity costs increase as the hospital becomes more specialized).
b. What is the opportunity cost of the hospital serving 4,000 rather than 2,000 in-patients yearly?
c. What is the opportunity cost of the hospital serving 21,000 rather than 12,000 out-patients yearly?

Exercise 1.4
Refer to production possibilities information in Exercise 1.3.

a. Explain why and how the hospital is considered productively inefficient by serving 4,000 in-patients and 20,000 out-patients yearly?
b. Which alternative, *B* or *C*, is allocatively efficient? What type of information would help you determine this?

Exercise 1.5
How would you go about deciding if a move from *B* to *C* were Pareto Optimal? Assume that the production possibilities in Exercise 1.3 referred to the nation as a whole rather than a single hospital.

Exercise 1.6
Refer to the production possibilities information in Exercise 1.3.

a. If the hospital's capacity, staff, and equipment increased, how would this affect your graph?
b. Suppose a new technique is discovered that reduces the amount of resources needed to process registration and billing for both types of patients. How would this affect your graph?

Exercise 1.7
Two countries, Scandia and United, are capable of producing output at the alternatives given in the table.

Scandia's possibilities *Value of output produced in the:*	*A*	*B*	*C*	*D*	*E*	*F*
Government sector	100	80	60	40	20	0
Private sectors	0	20	40	60	80	100

United's possibilities *Value of output produced in the:*	*G*	*H*	*I*	*J*	*K*	*L*
Government sector	300	240	180	120	60	0
Private sectors	0	60	120	180	240	300

a. Scandia is currently producing at alternative *B* and the United at alternative *K*. Are both countries achieving productive efficiency? Explain.

b. Scandia is currently producing at alternative *B* and United at alternative *K*. Are both countries necessarily achieving allocative efficiency? What assumption are you making about preferences in answering yes or no?

Exercise 1.8
Refer to the table in Exercise 1.7. If United were producing goods in the government sector worth 50 and goods in the private sectors worth 230, is it possible to increase its production of public goods without reducing the amount of goods produced in the private sectors? Explain.

Exercise 1.9
"Redistributing national healthcare expenses from a few terminally ill elderly to improve the public health of children is economically efficient." Discuss this statement in terms of the Pareto Criterion.

Exercise 1.10
Clients, as compared with customers, can be viewed as submitting themselves to a process that may or may not achieve desired outcomes. Are students and patients more like clients than customers? Explain.

Exercise 1.11
Give an example of a NFP firm providing a service for which clients pay a fee. What information do you need to estimate the full cost of providing this service? Do you suppose that the full cost is less or equal to the user fee? Explain.

CHAPTER 1 APPENDIX: REGRESSION TUTORIAL

In surveying this text, you find several Application employing a statistical method called regression. Researchers and practitioners in public and nonprofit administration use regression to test and understand relationships. After reviewing the statistical tools underlying this research, you should be able to understand and test relationships in the increasing amounts of data presently available to public managers.

The simplest forms of regression analysis involve just one independent variable. The dependent variable is usually designated Y, and the independent variable is represented with an X. The relationship or model we seek to find could then be expressed as:

$$Y = a + bX$$

This is called simple or bivariate linear regression (BLR) model because there are just two variables: Y and X.

In this expression, a represents the intercept or constant term for the regression equation. The intercept is where the regression line crosses the vertical, or Y, axis. Conceptually it is the value that the dependent variable (Y) would have if the independent variable (X) had a value of zero. We will have more to say about the interpretation of the intercept later.

The value of b represents the slope of the regression line. The slope is the rate of change in the dependent variable for each unit change in the independent

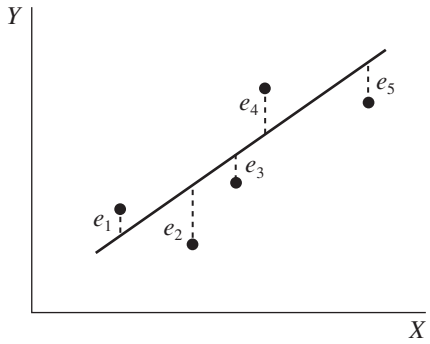

Figure A1.1 Ordinary least squares regression line for *Y* as a function of *X*. Residuals, or deviations, between each point and the regression line are labeled e_i.

variable. Understanding that the slope term (b) is the rate of change in Y as X changes will be helpful to you in interpreting regression results. If b has a positive value, Y increases when X increases and decreases when X decreases. On the other hand, if b is negative, Y changes in the opposite direction of changes in X.

The most commonly used criterion for the "best" regression line is that the sum of the squared vertical differences between the observed values and the estimated regression line be as small as possible. To illustrate this concept, Figure A1.1 shows five observations of the relationship between some Y variable and some X variable. You can see from the scattering of points that no straight line would go through all of the points. We would like to find the one line that fits closest to all the points; the regression line meets this criterion. It minimizes the sum of the squared vertical deviations of the actual observed values from the best-fit regression line.

The vertical distance between each point and the regression line is called a deviation. We will represent these deviations with e_i (where the subscript i refers to the number of the observation). A regression line is drawn through the points in Figure A1.1, and the deviations between the actual data points and the estimates you would make from the regression line are identified as e_1, e_2, e_3, and so on. Note that some of the deviations are positive (e_1, and e_4), while the others are negative (e_2, e_3, and e_5). Some errors are fairly large (such as e_2), while others are small (such as e_3). By our criterion, the best regression line is that which minimizes the sum of the squares of these deviations. The deviations from the regression line are also frequently called residuals. You are likely to see the term residuals used in printouts from computer programs that perform regression analysis.

The method of finding the values of a and b (that is, the regression line) that minimizes the sum of these squared errors is called ordinary least squares regression (OLS). Using the method of ordinary least squares, we square each of the deviations and add them up. We square the deviations so that positive and negative deviations do not cancel each other out as we find their sum. The single line that gives us the smallest sum of the squared deviations from the line is the best line according to the method of ordinary least squares.

A1.1 Interpreting the Intercept and the Slope of a Regression Model

How should we interpret the regression equation? In particular, what do the estimated values of (a) and slope (b) mean? The value of the intercept (a) indicates, at least conceptually, that if we were to extend the regression line to its intersection with the vertical axis (the point where the X variable is zero), the value of the dependent variable would be a.

We must be cautious in making such interpretations. In many cases, it is erroneous to interpret the value of a as the expected value of the dependent variable when the independent variable is zero because often our data do not include zero. In most cases the size of a is best interpreted as a positioning parameter for the height of the function. We really do not know what the function looks like as we go beyond the boundaries of our data. Recall that b represents the slope of the regression function. That is, the value of b tells use the rate of change in the dependent variable per unit change in the independent variable.

A1.2 Underlying Assumptions of the Ordinary Least Squares (OLS) Regression Model

Several assumptions underlie the ordinary least squares regression model. A general understanding of these assumptions is necessary to appreciate both the power and limitations of OLS regression:

1 For each value of X there is a conditional probability distribution Y. Figure A1.2 shows the conditional probability distributions of Y for two of the possible values of X (X_1 and X_2). Y is specified as the dependent variable and X as the independent variable. The means of the conditional probability distributions are assumed to lie on a straight line, according to the following equation: $Y = a + bX$. In other words, the mean value of the dependent variable is assumed to be a linear function of the independent variable (note that the regression line in Figure A1.2 is directly under the peaks of the conditional probability distributions for Y).

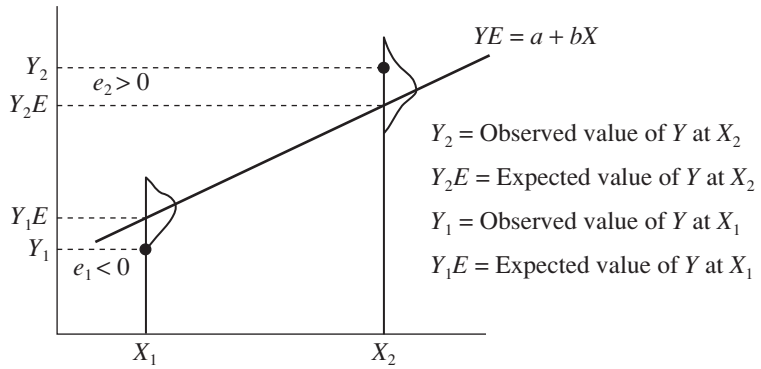

Figure A1.2 Distribution of Y values around the ordinary least squares regression line. For any X the possible values of Y are assumed to be distributed normally around the regression line. Further, the residuals (e) are assumed to be normally distributed with a mean of zero and a constant standard deviation.

2 OLS assumes that the standard deviation of each of the conditional probability distributions is the same for all values of the independent variable (such as X_1 and X_2). In Figure A1.2 the "spread" of both of the conditional probability distributions shown is the same (this characteristic of equal standard deviations is called homoscedasticity).
3 The values of the dependent variable (Y) are assumed to be independent of one another; so if one observation of Y lies below the mean of its conditional probability distribution, this does not imply that the next observation will also be below the mean (or anywhere else in particular).
4 All of the conditional probability distributions of the deviations or residuals are assumed to be normal. That is, the differences between the actual values of Y and the expected values (from the regression line) are normally distributed random variables with a mean of zero and a constant standard deviation.

These four assumptions may be viewed as the ideal to which we aspire in calculating a regression line; while these underlying assumptions of regression are sometimes violated in practice, they should be followed closely enough to ensure that estimated regression equations represent true relationships between variables. For the practitioner, it is important to note that if these four assumptions are not at least closely approximated, the resulting OLS regression analysis may be flawed. Summary statistics generally provided in most regression computer packages allow us to check compliance with these assumptions. These statistics are described below in this Appendix, as well as the likely outcomes of violating these assumptions.

A1.3 Evaluation of OLS Regression Models

In this section we describe a relatively simple process that will help you evaluate OLS regression models. This process is summarized by a set of questions to ask when evaluating regression models, either those you developed or others. These questions are:

1 Does the model make sense? (That is, is the model consistent with a logical view of the situation being investigated?)
2 Is there a statistically significant relationship between the dependent and independent variables?
3 What percentage of the variation in the dependent variable does the regression model explain?
4 Is there a problem of serial correlation among the error terms in the model?

Let us now consider each of these questions and how they can be answered.

Step 1: evaluate whether the model makes sense

Never use a model that does not make sense. If the results are at odds with what logic suggests something must be wrong. For example, suppose you look at a regression of a hospital's admissions (S) as a function of the fees (P) charged per day and see the following result:

$$S = 240 + 1.22P.$$

What would you think? Would you go to management and suggest that if they want to increase admissions they should raise fees? And, if the initial price increase does not increase admissions enough they should keep increasing price because admissions are positively related to the price that is charged. Certainly not!

Clearly something is wrong with this model. Economic logic (as well as considerable empirical evidence) supports the notion that admissions and fees are inversely related in nearly all cases. Thus, we expect that the sign for the slope would be negative. When the sign of the slope term is not what logic suggests, the model does not make sense and should not be used.

The situation discussed above is not an uncommon finding when regressing sales of any good or service on price. What then is the problem that leads to such an illogical result? In this case the problem is that the model is probably *under specified*, meaning that there are additional factors that have not been included in the model that have caused sales to go up despite price increases rather than because of price increases. For example, perhaps incomes have also increased, or the size of the market has expanded due to the closing of other hospitals, or greater advertising has increased product demand. Later when we discuss multiple regression we will see how such other factors can be included in a larger regression model.

There is no statistical test to determine whether or not the model makes sense. You must do that yourself based on your understanding of the relationship being modeled. If you cannot make the correct judgment about the appropriate sign you probably do not know enough about the area of investigation to be working with the model.

Step 2: check for statistical significance

Suppose you ask 20 people to each give you the first number that comes to their mind. Then you split the 20 numbers into two sets of 10, calling one set Y and the other set X. After entering these data into a computer regression program, regress Y on X, i.e., Y as a function of X. The program churns out an intercept and slope. But does the regression equation have any useful meaning? Would you expect to find a real functional relationship between Y and X? With these two sets of numbers, the answer is no to both questions. That is, if the values for X and Y are arbitrarily selected, you would not expect to find a functional relationship between them.

If Y is not a function of X, the best estimate of Y is the mean of Y (Y_M), regardless of the value of X, since Y does not depend on X. If this is the case, the regression line would have a slope equal to zero ($b = 0$). The scatter gram in Figure A1.3 illustrates such a case for the following data:

Y	14	19	10	14	9	18	20	13	8	19	14	16
X	5	6	7	10	13	13	13	15	18	18	21	21

Table A1.1 is the regression output produced from entering the above X and Y variables into data analysis under "Tools" in an Excel spreadsheet. The OLS regression

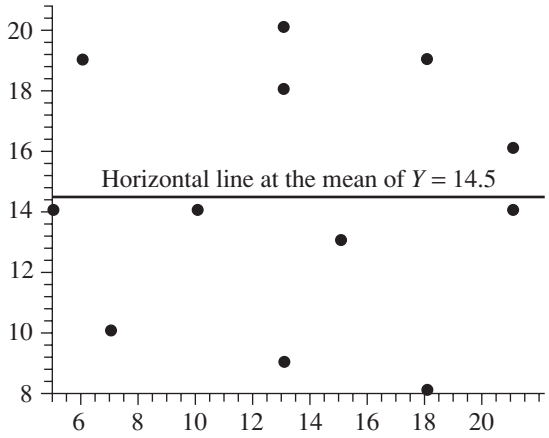

Figure A1.3 Scattergram when *Y* is not a function of *X*. Based on these 12 observations there does not appear to be a functional relationship between *X* and *Y*. When no such relationship exists, the best estimate of *Y* for any observed *X* is the mean of *Y* (Y_M).

equation for this set of points is: $Y = 14.66 - 0.012X$. The mean value of *Y* is 14.5. Note that the intercept value in the regression equation (14.66) is very close to the mean value of *Y*. In fact, if you draw the regression equation on Figure A1.3, you will find that it is very close to the horizontal line already drawn at the mean of *Y*.

Table A1.1 Regression results provided in the Excel spreadsheet

Summary output

Regression statistics	
Multiple *R*	0.016156
R square	0.000261
Adjusted *R* square	−0.09971
Standard error	4.253854
Observations	12

Anova

	df	*SS*	*MS*	F	*Significance* F
Regression	1	0.047244	0.047244	0.002611	0.960255
Residual	10	180.9528	18.09528		
Total	11	181			

	Coefficients	*Standard error*	t *Stat*	P-*value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	14.65748	3.317646	4.418037	0.001298	7.265305	22.04966	7.265305	22.04966
X	−0.01181	0.231151	−0.0511	0.960255	−0.52685	0.503226	−0.52685	0.503226

Residual output

Observation	*Predicted* Y	*Residuals*
1	14.59843	−0.59843
2	14.58661	4.413386
3	14.5748	−4.5748
4	14.53937	−0.53937
5	14.50394	−5.50394
6	14.50394	3.496063
7	14.50394	5.496063
8	14.48031	−1.48031
9	14.44488	−6.44488
10	14.44488	4.555118
11	14.40945	−0.40945
12	14.40945	1.590551

We need some method of evaluating regression equations to see if there is a meaningful functional relationship between Y and X. This is done by using a t-test to see if the estimated slope (b_E) is statistically significantly different from zero. If it is, there is sufficient evidence in the data to support the existence of a functional relationship between Y and X.

However, if the value of b_E is not significantly different from zero, we would conclude that Y is not a linear function of X.

A t-test is used to test the null hypothesis that the slope of the true relationship between Y and X is equal to zero. You could write this null hypothesis as

$$H_0: \beta = 0$$

The t-statistic used for this t-test is calculated as follows:

$$t_c = (b_E - 0) \div (\text{SE of } b_E)$$

Where SE of b_E is the standard error of b_E (the standard deviation of the probability distribution of the estimator). The standard error is included in the output of virtually all regression programs; it is derived from the residuals, or deviations, between each point and the regression line. The value of t_c indicates how many standard errors our estimate of β is from zero. The larger the absolute value of the t-ratio, the more confident you can be that the true value of β is not zero. Most regression programs provide the calculated t-statistic (t_c) as a standard part of regression output.

The statistical test of the significance for a regression coefficient can take any of the following three forms:

Case 1: H_1: $\beta \neq 0$

This form is appropriate when you are just testing for the existence of any linear functional relationship between Y and X. In this case, you have no reason to think that the slope will be either positive or negative.

Case 2: H_1: $\beta < 0$
This form is appropriate if you think that the relationship between *Y* and *X* is an inverse one. That is, you would use this form when you expect an increase (decrease) in *X* to cause a decrease (increase) in *Y*.

Case 3: H_1: $\beta > 0$
This form is appropriate if you think that the relationship between *Y* and *X* is a direct one. That is, you would use this form when you expect an increase (decrease) in *X* to cause an increase (decrease) in *Y*.

Situations such as described in case 1 imply the use of a two-tailed test. This means that if we want to have a 95 percent confidence level, we would have 2.5 percent of the total area under the *t*-distribution in the outer part of each tail of the *t*-distribution, as illustrated here in Figure A1.4.

A 95 percent confidence interval is the same as a 5 percent significance level. The confidence level and the significance level always sum to 1.0. The symbol α (alpha) is usually used to represent a significance level. Thus, in this example, $\alpha = 0.05$, and since we have a two-tailed test we split the significance level between the two tails of the distribution. That is, $\alpha/2$ is the area under each tail of the *t*-distribution.

In performing a *t*-test, we not only have to decide on a significance level, but we also have to correctly identify the number of degrees of freedom to use. In

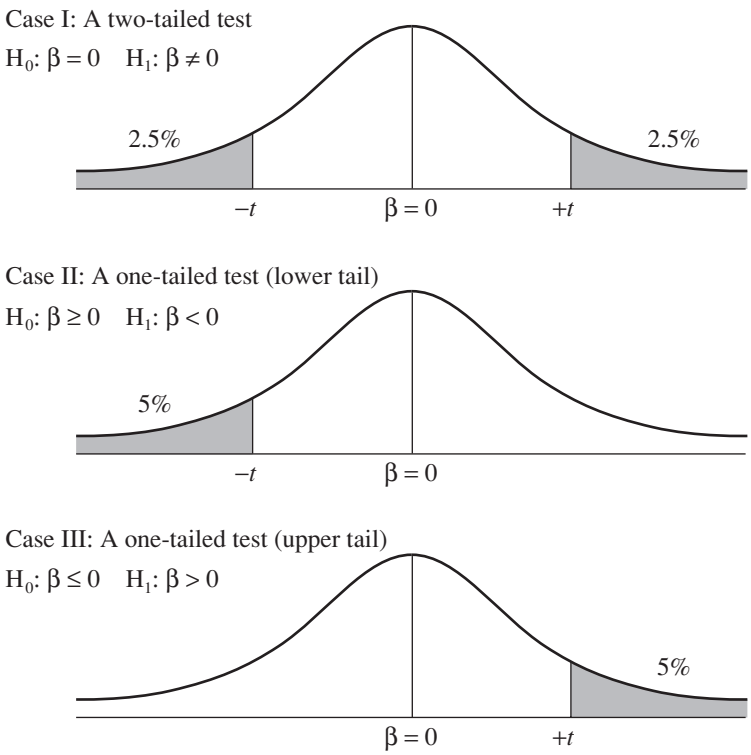

Figure A1.4 Cases 1, 2, and 3.

bivariate linear regressions, the appropriate number of degrees of freedom (*df*) is: $df = n - 2$. Where n equals the number of observations used in determining the values of a and b and 2 is the number of parameters estimated.

For a two-tailed test we reject the null hypothesis if the absolute value of t_c is greater than the table value of t at the desired confidence level. As an example of using a two-tailed *t*-test, consider the data shown in the scatter gram of Figure A1.3. In that case, you would have no reason to expect the slope to be either positive or negative. You would just be testing to see if the estimated value was indeed statistically significantly different from zero. Recall that the regression equation for that data is:

$$Y = 14.66 - 0.012X$$

Our null hypothesis is that the slope is equal to zero; that is, H_0: $\beta = 0$. The alternative hypothesis is that the slope is not equal to zero; that is, H_1: $\beta \neq 0$. The regression output on your calculator or computer should indicate that the standard error of b for this regression is 0.231, so we find t_c as follows:

$$t_c = (-0.012 - 0) \div (0.231) = -0.052$$

There were 12 observations, so the number of degrees of freedom is: $df = 12 - 2 = 10$. For a two tailed test with a 0.05 significance level, the critical absolute value (ignoring the sign) for t would have to be greater than 2.228, based on *t*-tables available in most statistical textbooks. A rule of thumb is often used in evaluating *t*-ratios when a *t*-table is not handy. The rule is that the slope term is likely to be significantly different from zero if the absolute value of the calculated *t*-ratio is greater than 2. This is a handy rule to remember as you analyze regression studies presented in the Application.

In this sample of 12 pairs of random numbers, the absolute value of the calculated t of -0.052 is not greater than 2, so we do not have enough evidence to reject the null hypothesis that the value of the coefficient is really 0. In fact, the *P*-value, provided in Table A1.1 indicates that if the true value of coefficient were 0, the statistical deviation from 0 in our example of random numbers is expected in approximately 96 percent of similar tests. Therefore, we conclude that there is no statistically significant linear relationship between Y and X at a 5 percent significance level ($\alpha = 0.05$).

Situations such as those described in cases 2 and 3 call for the use of a one-tailed test because we are only concerned with being either below zero (case 2) or above zero (case 3). In these cases, if we want to be 95 percent confident, we would have the entire 5 percent significance level in the outer part of either the lower (case 2) or upper (case 3) tail of the distribution. The number of degrees of freedom is still equal to $n - 2$.

Step 3: determine the explanatory power of the model

The dependent variable (Y) used in a regression analysis varies with the value of the independent variable (X). Otherwise there would be no reason to try to model Y. Therefore, it is convenient to have a measure of how much of that variation in Y is explained by the regression model. That is just what the coefficient of determination (R^2) does for us.

The coefficient of determination (R^2) tells us the percentage of the variation in the dependent variable (Y) that is explained by the regression model. The worst

possible explanatory power a model could have is to explain none of the variation in the dependent variable ($R^2 = 0$), and the best possible model would be one that explains all of the variation in the dependent variable ($R^2 = 1.0$).

The coefficient of determination (R^2) will always be a value between 0 and 1. The closer it is to 0 the lower the explanatory power of the model, while the closer R^2 is to 1 the greater is the explanatory power of the model. Suppose $R^2 = 0.67$ for a regression model. This would then mean that 67 percent of the variation in the dependent variable is explained by that model.

We rarely calculate the coefficient of determination by hand. Virtually all regression analyses are done using a computer. The coefficient of determination provided on a computer printout, is generally identified as "*R* squared" or "R^2." In our example in Table A1.1 of made-up pairs, *R* squared is 0.0003 percent; variations in *Y* are not explained well by variations in *X*.

Step 4: check the distribution of the error terms

One assumption of OLS regression is that the error, or residual, terms are normally distributed random variables with a mean of zero and a constant variance. Therefore, we do not expect to find any regular pattern in the error terms. Whenever a significant time pattern is found in the error terms, serial correlation is indicated.

Figure A1.5 illustrates the two cases of serial correlation. In the left-hand graph, negative serial correlation is apparent. Negative serial correlation exists when a negative error is followed by a positive error, then another negative error, and so on. The error terms alternate in signs. Positive serial correlation is shown in the right-hand graph in Figure A1.5. In positive serial correlation, positive errors tend to be followed by other positive errors, while negative errors are followed by other negative errors.

With serial correlation, problems occur in using and interpreting the OLS regression function. Serial correlation does not bias the estimated coefficients, but it does make the standard errors smaller than the true standard errors. The *t*-ratios calculated for each coefficient will be overstated, which in turn leads to rejecting null hypotheses that should not have been rejected. That is, regression coefficients may be deemed statistically significant when indeed they are not. In addition, the existence of serial correlation causes the R^2 to be unreliable in evaluating the overall significance of the regression function.

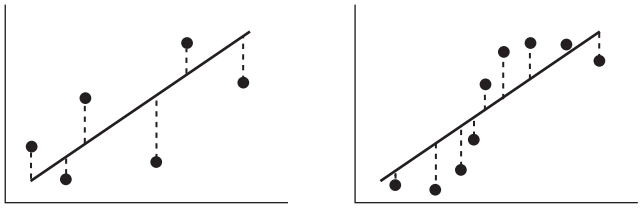

Figure A1.5 Negative and positive serial correlation problems. Negative serial correlation is illustrated on the left, positive on the right. The residuals are indicated by dashed lines.

There are ways to test for serial correlation. The method most frequently used is the Durbin–Watson statistic (DW). It is calculated as follows:

$$DW = \sum(e_t - e_{t-1})^2 \div \sum e_t^2$$

where e_t is the residual for time period t and e_{t-1} is the residual for the preceding $(t - 1)$ time period. Regression analysis in statistical software generally include the Durbin-Watson statistic, so you are not likely to ever have to calculate it directly. This statistic is not provided in Table A1.1. You may, however, check the residuals for any obvious patterns, although with merely 12 observations any observation or calculated DW (2.51 in this example) may be misleading.

The DW statistic will always be in the range of 0 to 4. As the value of the DW statistic approaches 4, the degree of negative serial correlation increases. As the value of DW approaches 0, positive serial correlation appears more severe. As a rule of thumb, a value close to 2 indicates that there is no serial correlation.

You might wonder what causes serial correlation. A primary cause of positive serial correlation, the most common form in business/economic analyzes, is the existence of long-term cycles and trends in the data. Serial correlation can also result from a misspecification of the model.

If we find a problem with serial correlation we can try several relatively simple things to reduce the problem. One is to use first differences of the variables rather than the actual values when performing the regression analysis. That is, use the change in each variable from period to period in the regression. Other potential solutions involve adding additional variables and/or nonlinear terms to the model.

A1.4 Point and Interval Estimates

Regression equations are often used to make estimates of the value of the dependent variable for a given value of the independent variable. When such estimates are made, it is common to give both a point and an interval estimate. The point estimate is generated directly from the regression equation.

Point estimates are useful. Estimating a precise number is often preferred to a range, but it is generally not accurate. Thus, it is often preferable to make an interval estimate in such a way that we can say we are 95 percent (or some other percent) confident that the true value will be somewhere in the interval. A simple approximation for a 95 percent confidence interval for a general bivariate regression model can be given as:

$$Y = Y_E \pm 2(\text{SEE})$$

where Y_E is the point estimate and SEE is the standard error of the estimate. The value for SEE is part of the output of nearly all regression programs.[1]

[1] However, it can also be easily calculated as follows:

$$\text{SEE} = [\,(\sum(Y_i - Y_{iE})^2 \div (n - 2)\,]^{0.5}$$

where n is the number of observations used in the estimation of the regression equation. (The 0.5 power is the same as the square root).

A1.5 Multiple Linear Regression

In many applications, the dependent variable is a function of more than one independent variable. In such cases, a form of OLS regression called multiple linear regression is appropriate. Application in this text contain several multi-linear regressions. This technique is a straightforward extension of simple linear regression and is built on the same basic set of assumptions.

The general form of the multiple linear regression model is:

$$Y = f(X_1, X_2, \ldots, X_n)$$

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n$$

where Y represents the dependent variable, and the X_i terms represents different independent variables. The intercept, or constant, term in the regression is α, and the β_i terms represent slope terms, or rates of change, for respective independent variables.

The addition of more independent variables to the basic regression model is helpful in developing better models. Doing so, however, adds to the four step evaluation process discussed previously. You will recall that those four steps involved answering these questions:

1 Does the model make sense? (That is, is the model consistent with a logical view of the situation being investigated?)
2 Is there a statistically significant relationship between the dependent and independent variables?
3 What percentage of the variation in the dependent variable does the regression model explain?
4 Is there a problem of serial correlation among the error terms in the model?

For multiple regression we need to add a fifth question.

5 Does there appear to be multicolinearity among the independent variables?

Multicolinearity

Multicolinearity results when the independent variables are highly correlated with one another. Whenever multicolinearity exists, the regression may not be reliable. In particular, coefficients may be incorrect.

Two factors might indicate a multicolinearity problem. First, if the standard errors of the coefficients are large relative to the estimated coefficients (resulting in unacceptably low *t*-ratios) for variables that you expect to be significant, multicolinearity is likely. Second, if pairs of independent variable have high correlation coefficients, a multicolinearity problem may exist. It is therefore important to examine the correlation coefficients for all pairs of the independent variables included in the regression. If two or more independent variables move together, their relationship introduces bias in explaining the dependent variable. One should avoid using pairs of independent variables that have simple correlation coefficients much above 0.7. In practice, with the data that we have in economic analyses, this is sometimes a high standard to live with and we may end up using pairs of variables that have a higher correlation if everything else in the model is acceptable.

The existence of multicolinearity may explain why a coefficient's sign is contrary to expectations, such as including both SAT (Scholastic Aptitude Test) scores and high school grade point average to determine a student's university grade point average. These two measures have a very high correlation. One would expect a positive relationship between high school and university grade point average, but if SAT results were to be included in the model it is quite likely that one would have a negative coefficient causing us to reject the model in the first step of our evaluation. Using two highly correlated variables as independent variables is an example of *over specification* of a regression model. So you now see that either under specification or over specification of a model may lead to signs for coefficients that are counter intuitive and thus would lead to rejection of the model on the basis of economic logic.

When multicolinearity exists, it does not necessarily mean that the regression function cannot be useful. The individual coefficients may not be reliable, but as a group they are likely to contain compensating errors. One may be too high, but another is likely to be too low (even to the extreme of having signs which are the opposite of your expectations). As a result, if your *only* interest is in using the regression for prediction, the entire function may perform satisfactorily. However, you could not use the model to evaluate the influence of individual independent variables on the dependent variable. Thus, one would rarely, if ever, use a model for which multicolinearity was a problem. Some things can be done to reduce multicolinearity problems, such as removing all but one of the highly intercorrelated variables from the regression.

The adjusted $\bar{R}^2$

When working with multiple regression models the adjusted coefficient of determination is used rather than the simple R^2 because the unadjusted R^2 will always increase as *any* new independent variable is added to a model, whether the variable is relevant or not. The adjusted coefficient of determination is usually designated R^2. It so happens that adding any additional independent variable will cause R^2 to go up but may or may not cause R^2 to rise. Thus, in interpreting multiple linear regression results one should always look at the adjusted coefficient of determination to evaluate the explanatory power of the model rather than the standard R^2.[2]

The diagrams in Figure A1.6 illustrate relatively high and low coefficients of determination for multiple linear regression functions. The graph at the top of the figure illustrates a regression plane with college applications as a function of tuition and a quality of education index. Six data points are shown to be not far from the

[2] The relationship between R^2 and $\bar{R}^2$ is:

$$\bar{R}^2 = 1 - (1 - R^2) \div [(n - 1)/\{n - (K + 1)\}]$$

where n represents the number of observations and K represents the number of independent variables. You can see that if n is large relative to K there will be little difference between R^2 and R^2.

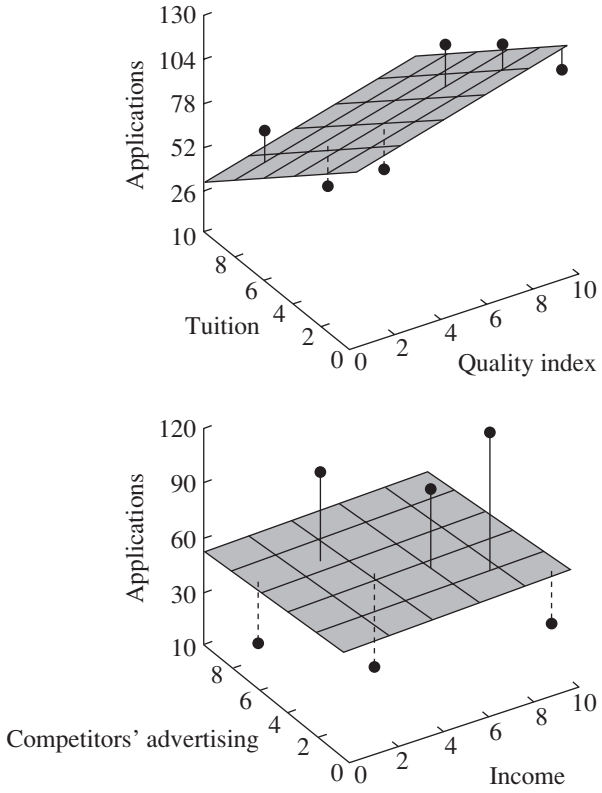

Figure A1.6 Two regression planes with different *R*-squares. The upper regression plane would have a higher coefficient of determination than would the lower one. Note that the six data points in the upper graph cluster closer to the plane than do the six data points in the lower graph.

regression plane and so the coefficient of determination would be relatively high. The lower graph shows a college's applications as a function of competitors' advertising and income. Here you see that the data points are further from the regression plane and so the coefficient of determination would be lower. Therefore, a model using tuition charges and perceived quality better explains the number of applications than a model using a competitor's advertising and income.