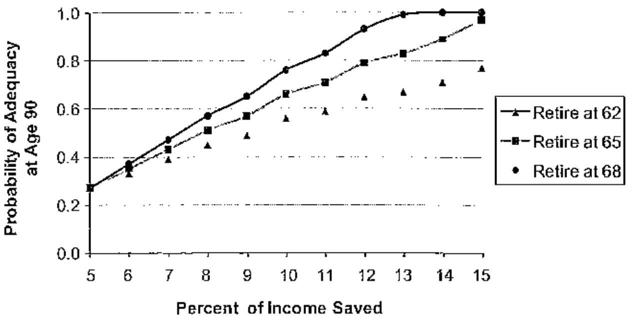

Figure 2.3. Probability of Adequacy as a function of Percent of Income Saved and Date of Retirement.

the same basic data are used in a different way to express a somewhat different point:

> Changing the base-case savings rate (10 percent) by one percentage point adds about $45,000 to your assets on the date you retire.

Another way you can explore for insights is to examine alternative outcome measures. The amount saved at retirement is one simple outcome measure, but it may not be the most useful measure for your friend. His is probably anxious to know whether a certain amount saved at retirement will be "adequate." But how do you measure adequacy? You could measure assets at death. You could measure years of comfortable retirement. You could estimate the probability of having assets left at age 90 or passing on at least $1M to heirs. These are just a few examples from an unlimited set of possibilities.

This exploration might lead to a result like Figure 2.3, which shows the probability of having adequate retirement savings through age 90 as a function of the percent of income saved and the age at retirement. This figure packs a lot of useful information into a small space in a form that is immediately clear to the client. (We explore this problem more fully in Chapter 5.)

To summarize:

- Insights are useful conclusions about the problem expressed in words.
- Insights are discovered through model exploration.
- Model exploration involves sensitivity analysis and analysis of alternative outcome measures.
- Generating insights requires a sound model, and the patience and skill to use the model to explore the problem from all angles until hidden gems emerge.

2.2 TOOLS FOR MODELING

We mentioned in Chapter 1 that modeling involves both technical and craft skills. Technical skills include, for example, knowledge of the Excel VLOOKUP function, the ability to formulate an optimization model, and the ability to interpret the results of a simulation model. Craft skills are more difficult to describe but no less essential to effective modeling. They include, for example, the ability to extract the essence of a problem, build a simple but effective model structure, and translate model results into useful insights.

Most craft skills are used in the background, out of sight. To teach them, we need to bring them into the light and to develop ways to talk about them. We have, therefore, identified six key tools for modeling ill-structured problems. We describe these tools in some detail here and show how they are used in practice throughout the book.

2.2.1 Influence Diagrams

Influence diagrams are a graphical means to display the structure of a model. They do not substitute for a quantitative model; rather, they provide a practical method for exploring alternative model structures during the early stages of problem framing. Influence diagrams are a far more effective tool than spreadsheets for creating a model structure. They encourage creativity, whereas spreadsheets encourage an inappropriate focus on details. In an influence diagram, the structure of a model is transparent; in a spreadsheet model, the structure is hidden behind the numbers.

Influence diagrams have just a few simple ingredients, each with its own symbol:

- An outcome measure (octagon)
- Intermediate variables (ellipse)
- Parameters (rectangle)
- Decision variables (parallelogram)

Consider this simple example from a well-structured situation. The problem is to determine what price to charge for a product to maximize profit. We start the influence diagram with the outcome measure, which is Profit (Figure 2.4). Next, we decompose profit into its constituent elements, which are Total Revenue and Total Cost (that is, Profit = Total Revenue – Total Cost). Total Cost, in turn, can be broken down into Fixed Cost and Total Variable Cost. Now, Total Revenue and Total Variable Cost are both a function of Quantity Sold, so we enter this variable in the diagram and connect it to both of the variables it influences. Total Variable Cost also depends on Unit Cost, so we include Unit Cost as a parameter at this point. Total Revenue is also influenced by Price, so this decision variable enters the diagram at this point. Price may

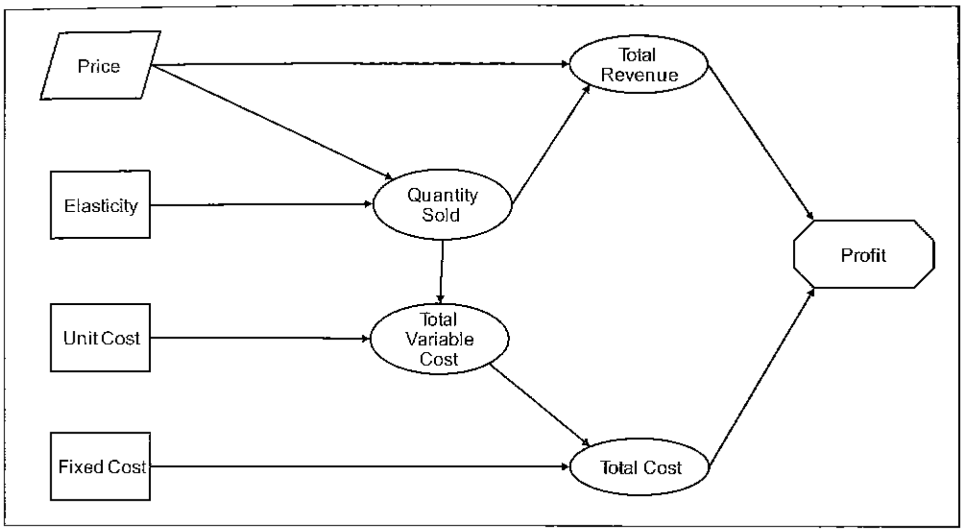

Figure 2.4. Influence diagram for Profit.

also influence Quantity Sold, so we connect Price to that variable and add another paramter, Elasticity, which influences Quantity Sold.

This diagram contains all four types of elements. The outcome measure, Profit, is shown enclosed in an octagon at the right of the diagram. The variables that influence Profit directly are some intermediate variables, shown as ellipses. Unit Cost, Elasticity, and Fixed Cost are parameters, shown as rectangles. Finally, the decision variable, Price, is shown as a parallelogram. Note how the influence diagram relates the decision variable Price to the outcome, Profit.

Here is another short problem statement, one that describes an ill-structured situation for which a model is needed.

Northlands Museum Capital Campaign

Northlands Museum is a regional museum of human and natural history located in a small city in a northern state. It has an annual budget of around $900K and an endowment of $1.2M. In recent years, the museum's board has reluctantly allowed the director to cover an operating deficit by transferring money from the endowment. The deficit has averaged $75K per year.

The museum director has recently proposed a major expansion of the museum, including building additional storage space for artifacts, upgrading exhibits, and installing air conditioning and humidity-control systems. A large capital campaign—the largest in the museum's history—will be required to fund this project. The board is concerned that the time and costs involved in the capital campaign will lead to additional requests to draw down the endowment.

The board has hired you to develop a model for analyzing the financial interrelationships among the operating budget, the capital campaign, and the endowment. The board's ultimate concern is that the endowment should not be smaller than it is now, when the campaign and the expansion project are completed in 10 or 15 years.

STOP **To the Reader:** Take a few minutes to draft an influence diagram for this problem before reading on.

Our influence diagram for this problem is shown in Figure 2.5. It shows the Endowment influenced by both Interest and Deductions. Deductions, in turn, depend on the Operating Deficit and the Capital Deficit (we are assuming that surpluses in either category also flow to the endowment). The Capital Deficit itself depends on Contributions to the capital campaign as well as on the cost of the campaign (Campaign Expenses) and on the costs of the proposed expansion (Expansion Costs). Operating Deficit depends on Operating Costs and Operating Revenue.

Although you need to practice to learn how to use influence diagrams effectively, here are a few principles to follow:

- Start by defining the outcome measure.
- Decompose the outcome measure into a small set of variables that determine it directly.
- Continue to decompose each variable into its constituents, drawing each variable only once.

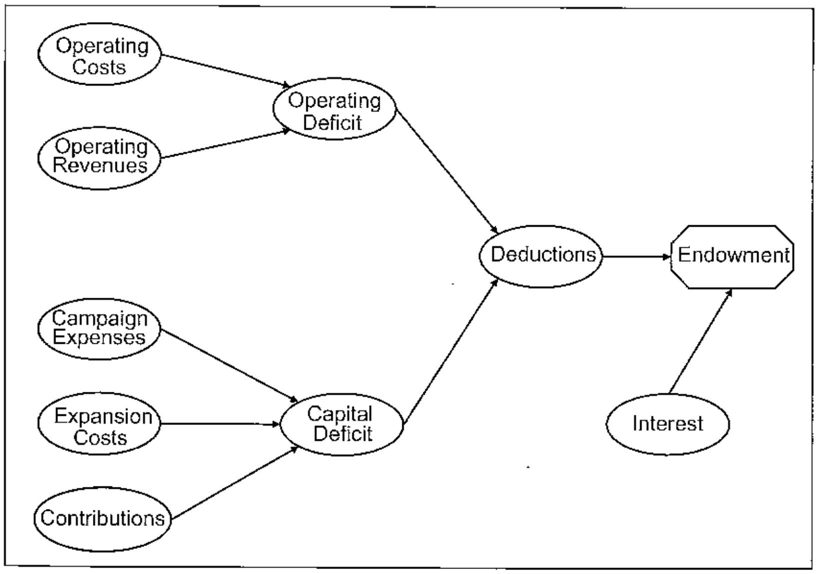

Figure 2.5. Influence diagram for Endowment.

- Link each variable to the variables it influences and is influenced by. (Variable A *influences* Variable B if Variable A is needed to *calculate* Variable B.)
- Identify parameters and decision variables.

2.2.2 Spreadsheet Engineering

Spreadsheets are the essential technical tools of business analysis. Effective modelers see to it that their spreadsheets are well structured and easy to use. Poorly structured spreadsheets not only are difficult to understand and work with, but they also are more likely to hide errors than are well-designed ones. (For more information on the risks of poorly designed spreadsheets, see the Spreadsheet Engineering Research Project website: http://mba.tuck.dartmouth.edu/spreadsheet/.)

Using spreadsheets effectively for modeling involves a set of skills we call *spreadsheet engineering*. Although problem framing is a highly creative and open-ended process, spreadsheet engineering (as the name implies) is more structured and routinized. The goal of spreadsheet engineering is to build models both *efficiently* and *effectively*. An efficient process leads to a completed model with minimum effort and waste; an effective process leads to an insightful analysis of the problem itself. Thus, it is necessary but not sufficient to build models efficiently; not doing so costs time and money. But it is still possible to build models efficiently and not be an effective modeler. Effective modelers use models to help their clients select and implement effective solutions.

We divide the process of spreadsheet engineering into four phases:

- Design
- Build
- Test
- Analyze

The essential idea in the design phase is actually to *design* your models, not simply to let them grow organically. A well-designed spreadsheet should have an easily understood logical and physical structure, with clearly defined modules. You should be able to distinguish data inputs from relationships and to document all essential assumptions.

Building a spreadsheet effectively requires you to have and use a sound design. In an effective engineering process, building the spreadsheet is mostly a mechanical task. You must pay attention to executing the design carefully without introducing flaws (or "bugs") into the model, rather than redesigning the spreadsheet. An effective spreadsheet builder knows where most bugs originate and can therefore take special care when it is required.

Model *testing* is a crucial and often overlooked phase in spreadsheet engineering. Because all spreadsheets develop bugs as they are built—many spreadsheets in use contain bugs, in fact—you need to develop a healthy skepticism about the presence of bugs in your own work. There are many effective ways you can test a spreadsheet for bugs. Some methods are built into Excel itself (although few analysts know about them). Other methods require special add-in software. Get into the habit of thoroughly testing all your models before using them for analysis.

The final phase in spreadsheet engineering is the *analysis* phase, where the payoff to your efforts is realized. Only a well-designed, well-built, and well-tested spreadsheet should be used for analysis. The analysis phase is where you can discover insights about the problem. Analysis is one of the most creative phases of modeling, but it too has its routine aspects. For example, there are several questions you should always ask, such as how sensitive the outcome is to the input parameters and which parameters have the biggest impact on the outcome? In many cases, you can answer these questions using specific Excel tools.

We present the principles of spreadsheet engineering in more detail in Chapter 3.

2.2.3 Parameterization

The basic concept of parameterization involves separating data inputs, or parameters, from the relationships that depend on those inputs. In other words, enter an input value into the model only once and use that input in all calculations that require it. This is essential to good spreadsheet modeling, because hiding numbers in formulas leads to confusion and errors. But the concept of parameterization is broader than this.

Parameterization in the broader sense involves choosing how to represent relationships in the most effective manner. By "most effective," we mean most effective in the analysis phase of modeling. Varying parameters to see how the output changes is an essential part of analysis. But you can only vary the parameters you have chosen to include in the model. Thus, the seemingly technical question of how to parameterize a relationship can have critical implications for how useful the model is for analysis. The example below will clarify what we mean.

Imagine you are trying to model the growth in market share for a new product. You could assume some initial market share and a constant annual growth rate, as in Figure 2.6a. Or you could assume share grows steadily from time T_1 until T_2, remains level at a given steady state level S until T_3, and then declines to zero at T_4 (see Figure 2.6b).

Our point here is not that one or the other of these relationships is better, but that each supports only certain kinds of analysis. For example, in the first model, you can vary the annual rate of growth of share but you cannot vary the steady state share. Likewise, in the second model, you can vary the time