1 The power and ubiquitousness of models

Consider the following problem: At exactly 6:00 AM, a monk leaves the base of a mountain to begin climbing the path to the top. He arrives at the top at 6:00 PM, where he spends the night in the monastery. The next morning at 6:00 AM, he begins descending along the same trail that he used to climb the mountain. At 6:00 PM he arrives back at the base of the mountain.

The question we pose is the following: Must he have been at some identical point on the trail at the exact same time of day on both his trip up and his trip down?

Take a few minutes to think about this problem before reading further and finding the answer. Be clear about what the problem asks: is there a spot on the trail, say, for example, at elevation 300 meters, that the monk reaches at the same time of day, for example 11:30 AM, both on his way up and his way down? You are not told whether or not the monk travels at a constant rate of speed or whether he rests sporadically during his journey. All you know is his departure and arrival times and that he travels the same trail. You are not asked to describe a particular time and place, only to decide whether there must be such a time and place.

This is a hard problem. When students are broken into a dozen or so small groups of four or five people and given ten minutes or so to work on this problem, most of the groups cannot decide if the answer is true or false. A few groups conclude it is false because it seems improbable to them. A few groups answer a slightly different question than the one asked. They make the following conditional claim: "If the monk travels at the same rate both up and down the mountain, he will be at the midpoint of the trail at noon on both days." They stop at this point because any enrichment of possibilities, e.g., the monk travels faster downhill for the first six hours, and then rests for awhile, makes it too difficult for them to figure out what might happen. Usually, only one or two groups conclude it is true. This is the correct answer.

How did the few groups who got the right answer arrive at their conclusion? Invariably in one of two ways. The most frequent method is where the group (or someone in the group) hits upon the idea of drawing a graph of the monk's progress. On the vertical axis they measure the height of the mountain, and on the horizontal axis they measure time. They then plot how high up the mountain the monk is at every moment of his twelve-hour trip up the mountain. Then they plot how high up the mountain he is at every moment of his twelve-hour trip down the mountain. If you try this, you will see that you cannot plot these paths without them crossing. The crossing point represents a point where the monk is at the same point on the path at the same time of day. A sample graph is depicted in Figure 3.1., with one hypothetical trip down and two different hypothetical trips up. The two trips up the mountain are drawn as solid lines and the trip down is drawn as a red line. One trip up the mountain has the monk going faster at first and slower at the end than does the other trip up the mountain.

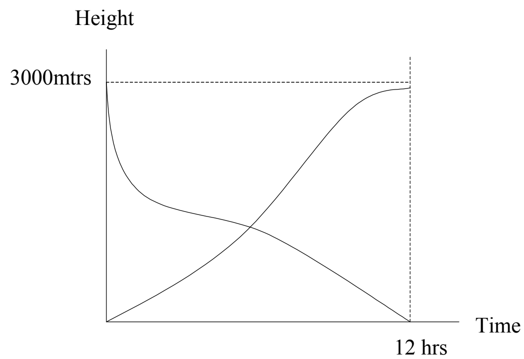

Figure 1:

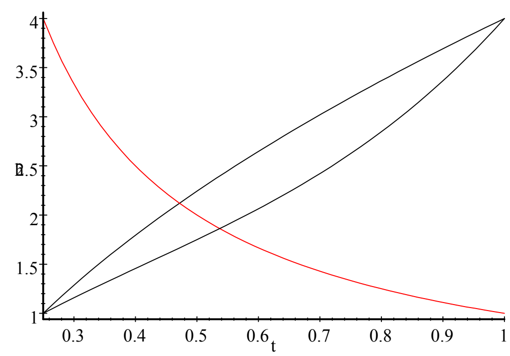

Figure 3.1: The Monk Problem

When this solution is presented to the other groups, most students immediately see that this is a correct analysis. However, a few students sometimes remain unconvinced. For them, the other method used to uncover the correct answer is the one that is persuasive.

This method is an argument by analogy. Invariably some group poses the following "thought experiment": Suppose there are two monks, one at the top of the mountain, one at the bottom. They both start to travel at 6:00 AM, one going down and one going up, and they both must finish their trip by 6:00 PM. Clearly they must meet on the path. If one thinks of the monk coming down the mountain as mimicking the pace of the lone monk of the first problem as he (the lone monk) descends on the second day, then the point where the two monks meet satisfies the requirement that the lone monk be at the same spot at the same time both on the way up and on the way down.

What this group has seen is that the two-monk thought experiment is essentially the same as the original problem. It is analogous or alike "enough" to the first problem, so we see that they capture the same phenomenon.

With both successful approaches, what students are doing to solve this problem is constructing a model: a logical representation of the essence of the situation. Note that this model is not an exact recreation of the problem: neither a graph or a two-monk thought experiment is *exactly* like the original problem. In fact, one can (with great difficulty) imagine a true skeptic who is not convinced by either of these models. Such a skeptic might want to follow the monk up and down the mountain, measuring distance and recording times. Of course, even then, all the skeptic has proven is that one particular monk on two particular days has been on the identical spot on the trail at the same time. To understand anything useful, that is, something that applies to other similar experiences, the skeptic still must abstract.

The point here is that to solve problems people construct abstract models. What a model does is eliminate insignificant or inconsequential detail, leaving the essence of the problem exposed for analysis. When economists tackle problems, this is what they must also do.

Models are not only essential for solving hard problems, they are also ubiquitous in everyday life. Consider what is meant by "a straight line". Most people remember the definition they learned in high school geometry that a straight line is the shortest distance between two points. They also generally agree that in everyday usage, many objects, such as the edge of a door, are straight lines. Upon reflection, though, everyone realizes that no object in the world satisfies the geometric definition: under a powerful enough microscope, every straight line will appear curved and jagged. We must conclude, then, that straight lines in the real world are to us like pornography was to Justice Potter Stewart of the Supreme Court: we can't define it, but we know it when we see it.

The conclusion we can draw from this discussion is that any attempt to make sense of the world involves abstraction and simplification. To make progress, we are forced to do things like some people do to solve the monk problem: "*Assume* there are two monks".

The reader may have heard the old joke about the shipwreck that left a chemist, a physicist, and an economist stranded on a deserted island. The only thing saved from the ship was a box of canned food. The three survivors' problem was how to open the cans. The chemist suggested they gather some roots and vegetation from the island from which he would make an acidic solution that they could use to eat through the top of the cans. The physicist suggested they use his eyeglass lens to focus sunlight on the cans and melt the top. The economist chuckled at the complexity of these two solutions, and then offered his own solution: assume a can-opener.

Even economists, despite being dismal scientists, can laugh at this joke. They know that they make many assumptions that appear heroic in their distance from what appears to be the real world. Again and again, students and non-economists express amazement at the plethora of assumptions used by economists, assumptions like: firms maximize profits, households choose a most-preferred bundle of commodities, and so on. The question one must keep in mind is not whether these assumptions exactly capture reality, but rather whether they capture the essence of the problem at hand, and thus help make progress in understanding or solving some economic problem.

In contrast to many other social science disciplines, economics has a rather well-developed concept of what are they key components of their models. These key components form a generic structure within which virtually all economic models fit. Once one has learned this generic structure, he or she can better remember and use the more specific models designed to address specific problems. For this reason, we are going to devote considerable attention to this generic structure. In the remainder of this chapter, we will develop this generic structure and illustrate it with concrete examples of a model with which the student might have some familiarity from a principles course.The structure, language, and depiction of economic models.

Like students working on the monk problem, economists work on economic problems. For example, at the end of World War I, governments wanted to reset exchange rates (the price of one currency in terms of another, e.g., dollars/pound) at fixed levels, after having been forced to let private market forces determine them during the war years. The problem was at what level they should be reset: if set at pre-war levels, many officials worried there might be dangerous deflation, unemployment, and severe trade imbalances. Work on this problem by the economist Gustav Cassel led to the Purchasing Power Parity Theory of exchange rate determination.

More familiar to the reader who has had an economics-principles course are two other problems, one from microeconomics and one from macroeconomics. An archetypal microeconomics problem is the determination of prices and of quantities of particular goods and services produced and consumed per unit of time. A more specific version of this question might be something like: What will happen to the price of wheat and to the quantity of wheat produced and consumed per year if Asian economies return to their pre-1997 high growth rate of GDP? A familiar macroeconomic problem is the determination of national income and output "in the short run". To be specific, one might ask, "What will happen to U.S. GDP over the next year if congress immediately raises government spending?

As discussed in the preceding section, understanding the world and working on any problem requires using models. In this section, we will explicitly delineate the "thought protocols" used by economists as they construct models. First we provide a description of the elements of models, and then we illustrate these concepts with a specific example. As we do this, we also introduce some basic terminology and review the key concepts from analytic geometry that are most useful to economists.

1.1 Elements of models

1.1.1 Variables

First, for the problem at hand, economists think about what variables are likely important and what variables are likely unimportant. A variable is a quantity free to take on any of a number of permissible values. For example, when thinking about what determines an individual's total expenditure on consumption per year (itself a variable), economists believe that variables like disposable income per year, wealth, age and perhaps interest rates are important. In contrast, they don't believe variables like hair color, height, phase of the moon, or sex are important. They may be important for the study of another problem, such as the proportion of expenditure spent on clothing, but not for the problem of explaining total expenditure. An economist makes this choice of what variables are important and what variables are inconsequential on the basis of a variety of things, including intuition, introspection, information gleaned from related disciplines such as psychology or sociology, and from empirical observations. This choice, never irrevocable, represents the first part of the art of modeling.

An important feature of variables used in economic modeling is that they can all be represented by numbers. This is really just a way of renaming things. Even things like phases of the moon or hair coloring can be assigned different numerical values For example, if hair color was thought to be important for some model, red could be assigned (i.e.,named) the number one (1), brown the number two (2), purple the number three (3), and so on. By assigning variables values that are numbers, economists are able to use the techniques of analytic geometry to depict these variables and their interactions with each other.

For the measurement of many variables, we abstract from their essential discreteness and approximate them as if they can be measured by *real numbers.* Real numbers are all those numbers which may be represented by terminating, e.g., $1\frac{1}{10}$ =1.1, or nonterminating, e.g., $\pi = 2.1417...$, decimals. The "dot-dot-dot" symbol after the digit "7" in our expression for π means that an infinite number of other numbers follow. For example, for purposes of development of a theory we may allow the 'permissible values' of things like consumption per year and income per year to be the set of non-negative real numbers. This is an approximation to reality because we can't really sub-divide the measurement of things like consumption per year and income per year more finely than small discrete units such as $\frac{1}{100}$ of a dollar per year. That is, we can't have a measurement of these variables that is, for example, the familiar transcendental number $\pi \approx 2.1417$. This lack of realism is offset by the convenience of thinking about the permissible values of such variables as coming from the *continuum* of values represented by the real line. The point here is not that we think most economic variables are measured continuously. In fact, we think most are measured in discrete units. For most models, though, these discrete units are quite fine, and consequently a listing of all possible measurements would be quite large. By assuming that economic measurements are members of the set of real numbers, we avail ourselves of the convenience of concise expressions and depictions of members of collections of variables. These collections of variables are important components of our models.

1.1.2 Logical structuring and representation

After making this first-stage decision, the economist then makes a logical structuring and representation of what he or she thinks are the interrelationships between these variables. There are two key parts to this process.

Endogenous versus exogenous variables First, the economist makes a further distinction among variables: he or she dichotomizes them into endogenous and exogenous. An exogenous variable is defined as having a value determined outside the model. That is, the value of an exogenous variable is taken by the economist as "god-given" and not to be determined by the economist. An exogenous variable is sometimes referred to as an independent variable. An endogenous variable, on the other hand, is defined as having its value determined jointly by the particular values taken by the exogenous variables and by the logical relationships between variables within the model. Endogenous variables are sometimes referred to as dependent variables.

Equations Once a decision is made about what variables are important and, of these, which variables are endogenous and which are exogenous, then the interactions between these variables must be specified. Usually these interactions are represented as a system of equations that consist of definitions, identities (things true by definition), technical descriptions, behavioral hypotheses, and equilibrium conditions. These relationships consist of statements like: "Consumption per year is an increasing function of disposable income per year", or "the quantity supplied per unit of time is an increasing function of the own price of the good".

There are a number of ways to depict a relationship between variables. Before describing these, let us be clear about what we mean by "relationship." When we state that 'quantity supplied per unit of time is an increasing function of the own price' we are asserting a *systematic* relationship between two variables: price, and quantity per unit of time. For a specific relationship, there are three ways to represent it: with a table of values, with a mathematical expression, and with a graph. A table of values is the most concrete way, but also the most cumbersome. A mathematical expression uses symbols to represent variables and algebraic operations to describe the relationships between these variables, and is both economical and flexible. Unfortunately, these very qualities make them more abstract and intimidating to many people.

Fortunately, because variables take on values that are numbers, many equations can be depicted as graphs. For many people, a visual depiction of mathematical statements provides greater understanding of the underlying logical relationships. This observation accounts for the long-standing appeal of analytic geometry as a tool for helping people understand abstract mathematics.

Most of the models used in undergraduate economics are simplified so as to permit their representation and manipulation with the tools of analytic geometry. As we work through our specific example in this chapter, we will review those tools most useful to us.

Parameters In some models, the equational statements tie variables together via variable-like entities called *parameters.* Parameters are like exogenous variables in that they represent a number, e.g., "six," and in that their values are given exogenously, that is, from outside the model.Solving the model

1.1.3 The canonical question

At this point, the model is complete. Now the economist frequently uses the model to answer the following canonical (conforming to a general rule; reduced to the simplest or most clear schema possible) question: What is the relationship between the values of the endogenous variables and the values of the exogenous variables? Answering this question is referred to as "solving the model."

The term "relationship" is not very precise, and needs some elaboration. For some models, usually ones that we develop as learning exercises rather than ones that we view as true working models, the form of this relationship is very exact. We want to know: What are the actual numerical values of the endogenous variables for any particular given permissible values of the exogenous variables? An extremely simple example of a physical model might make this concrete.

Consider a model for volume of a cube. Let volume be symbolized by the letter v, and the length, width, and height of a cube be symbolized by l, w, and h, respectively. Volume is the endogenous variable, and length, width and height are the exogenous variables. Our theory is that volume is the product of the length, width, and height of a cube. This would be expressed in symbolic mathematical form as the equation:

$$v = lwh.$$

For most of us, this hardly seems like a model, because we think of it as an accurate description of reality. Nonetheless, it serves to illustrate a particular type of relationship between endogenous and exogenous variables. For given permissible numerical values of the exogenous variables l, w, and h, the relationship between endogenous and exogenous variables gives us a numerical value for volume. For example, if $l = 2$, $w = 4$, and $h = 6$, then $v = 48$.

For many economic models, the specific form of the canonical question asked is a slight variation of the one posed above. It asks: What are the *qualitative changes* in the values of the endogenous variables for given, arbitrary *qualitative changes* in the values of the exogenous variables? That is, instead of specifying unique numbers as values for each of the exogenous variables and asking what are the associated numerical values of the endogenous variables, it posits a numerically-unspecified directional change, i.e., increase or decrease, in the value of an exogenous variable and asks: What are the associated numerically-unspecified directional changes in some or all of the endogenous variables?

The basic reason we ask a model this question about *changes* is that our economic models don't give us enough information to make the detailed specifications necessary to answer the first question. Many of the behavioral equations that make up our models can only be specified in the weaker terminology of qualitative changes. For example, a complete model of the interactions between an assumed exogenous variable such as the Fed Funds interest rate and other endogenous macrovariables such as employment and inflation might have as one equation: "Consumption is an increasing function of disposable income." There is information here: if disposable income goes up, consumption goes up. But the information not here is the quantitative relationship that tell us how much consumption goes up in response to a particular numerically-valued change in disposable income.

Without quantitative information embedded in our models, the best we can hope for is that our models can answer this form of the canonical question. The ability of many economic models to map values of exogenous variables to values of endogenous variables has been compared to the ability of a Russian tank commander in World War II to predict the response of his tank to a movement in the driving controls. These tanks were allegedly crude affairs, and movement of the tank was controlled by a big lever. If the lever was pushed forward, the tank moved forward. The farther forward the lever was pushed, the faster the tank would go. If the lever was pulled back, the tank went backward. The farther back the lever was pulled, the faster in reverse the tank went. The problem was that the lever was very "stiff:" the driver had to apply a lot of pressure to get it to move at all, and sometimes it would then move a lot, and sometimes a little. Hence, the commander could always tell in which direction the tank would move-forward or backward-but not how fast. Many economic models have that same characteristic: when an exogenous variable changes, the model can tell us in what direction the endogenous variables will move, but not by how much.

Why is this general question of the relationship between endogenous and exogenous variables the *raison d'etre* (reason or justification for existence) of an economic model? First, many of the real-world problems of interest to economists are problems for which policy advice is relevant and important. For example, the chairman of the Federal Reserve Board may want to know what will be the likely effects on national output, the inflation rate, the unemployment rate, and the exchange rate if he persuades the Fed to lower the interest rate it charges for loans to commercial banks. This interest rate is something the Fed can control if it so desires, so it can be treated as exogenous to the rest of the economy. In most macroeconomic models, a change in this exogenous variable engenders changes in the endogenous variables such as output, unemployment, inflation, and the exchange rate. Hence, this canonical question that economists ask of their models is exactly the question policy-makers should be interested in.

Second, this question helps economists avoid the particularly pervasive problem of confusing correlation with causality. As an example of this problem, consider the following theory: drinking diet cola makes people overweight. The evidence in support of this theory is the observation that one sees mostly overweight people buying diet cola at the grocery store. Another "theory" about another phenomenon we could propose is that the installation of storm windows brings on winter. The evidence in support is that every year the onset of winter is preceded by many people putting up their storm windows.

These two "theories" are relatively easily dismissed despite the presence of corroborating evidence that takes the form of correlation between the hypothesized cause and the effect. The dismissal is easy because we have good theories about what is exogenous and endogenous in these two examples. In both cases, the "evidence" is a correlation between the endogenous variables. What we really believe is that other latent (present though not directly visible) exogenous variables are changing, which are in turn changing the values of the endogenous variables. In the "diet cola leads to obesity" theory, we believe the truth is that already-heavy people are choosing diet cola in an attempt to reduce their caloric intake and get thinner. The latent variable that causes people to show up at the diet cola section and that causes these people to be heavy is "desire to get thinner." In the storm window example, the latent variable that causes both cold weather and people to put up their storm windows is the movement of the earth around the sun.

While in these two examples the fallacy inherent in each theory is transparent, political candidates, government officials and other public intellectuals often make similar errors. For example, one occasionally hears the argument that, because high school graduates make more money, are less likely to end up in jail, and in general have less troublesome life experiences, every public effort should be made to keep children in school until they graduate. Implicitly, anyone who makes this argument views high school attendance as an exogenous variable that directly affects endogenous variables such as lifetime earnings. Once this point is made clearly, other possibilities for which high school attendance is endogenous quickly spring to mind: perhaps high school drop-outs have done so poorly in previous schooling that they understand their best job prospects are in unskilled work.

1.1.4 Solution strategies

Sub-models Many economic models, in their entirety, involve many equations and many variables. As a strategy for understanding and explicating such a complex system, economists build sub-models that they then use to build the complete model. In these sub-models, variables (and sometimes parameters) that are ultimately and fundamentally endogenous within the model as a whole, are treated as exogenous.

Mathematics Powerful mathematical techniques are available for solving systems of equations, and an economic model is a system of equations. Advanced treatments of economics and economic research being described for other professional economists often uses these techniques. Even in these treatments, though, economists often use graphical techniques to help themselves and their

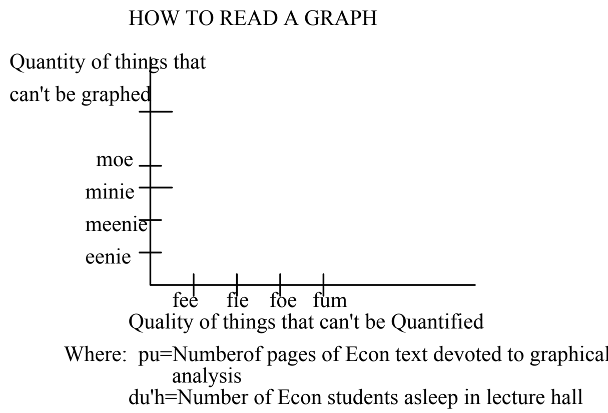

Figure 2:

readers understand the components of their models and how these components fit together. For non-economists, these graphical techniques are paramount for helping them understand a model.

Graphs A picture is worth a thousand words, and is also worth a few equations. Most people, including most economists, find graphs enormously helpful in understanding complex relationships between variables. Because variables take on values that are numbers, the techniques of analytic geometry can be used to depict economic models. For many people, though, the use of graphs is only slightly less confusing than equations in helping them understand an economic model.. For them, the following graph used by P.J. O'Rourke in his humorous layman's guide to economics, *Eat the Rich*, sums up their feelings about graphs in economics.

INSERT

Fortunately, there are some general features of how economists use these tools that will help people keep from getting lost in a thicket of complicated-looking graphs. First, it helps to know what visual constraints economists impose upon themselves and how these constraints shape their graphs Second, it also helps to know the final destination of the trip when the trip takes us through a maze of graphs.

Dimensionality Economic models, especially those used in international economics, frequently have many variables and many logical interrelationships between these variables. Unfortunately, a simultaneous depiction that human beings can understand of all these variables and interactions is impossible: most of us can't envision things in more than two, and at most sometimes three, dimensions. Hence, economists try to collapse information by a variety of tricks into two-dimensional pictures. This is frequently the sole reason for some seemingly-complicated graph; it is simply a two-dimensional way-station helping us to get from one part of a model to another. It is closely related to the use of sub-models.

The canonical question and economic graphs Remember that economists are trying to use their models to answer a particular question: what is the relationship between exogenous and endogenous variables? The final graphs in a presentation of a model (or sub-model, for that matter) are designed to answer this question. The strategy, then, is to depict on a two-dimensional graph two independent relationships between at most two endogenous variables. The intersection of these relationships determines a pair of numbers that is at least part of the solution to the model. We say this pair may be just a part of the solution because in models with more than two endogenous variable, there will have to be other numbers that are part of the solution, perhaps depicted in other related two-dimensional graphs. By judicious choices, the economist hopes to be able to have only one of these relationships shift in the two-dimensional plane when an exogenous variable changes its value. This allows one to "read off" from the graph the induced change in the values of the endogenous variables.

In the next section, we demonstrate these solution strategies in the context of one of the simpler economic models.Example

We now move on to an in-depth description of the usual ways that economists create and express the logical interactions between the variables in a model. To make the concepts concrete and memorable, we will illustrate them with the a specific model familiar to a student who has had an economic principles class: the microeconomic partial-equilibrium model of supply and demand for a particular good or service.

1.2 The microeconomic partial-equilibrium model of competitive demand and supply

This model, familiar from principles courses, is constructed from two sub-models: a model of demand, that is, a model of how individuals make decisions about how much of a particular good to consume over a given period of time; and a model of supply, that is, a model of how firms choose how much of the good to produce and sell during the given time period. The two sub-models are linked together by the equilibrium condition that demand equals supply. The endogenous variables for the model as a whole are the quantity per unit of time produced and consumed, and the equilibrium price. Exogenous variables are prices of other goods, income of consumers, and prices of inputs into the production process.

We use the modifier *partial equilibrium* because we are looking at interactions in one market only. This means we ignore potentially important feedback effects between one market and others. We are also deferring to a later point a full development of a deeper model of consumer and producer behavior. Our primary goal here is to provide a self-conscious description of a familiar model in terms of its generic parts.

1.2.1 Demand submodel

Variables First consider the demand sub-model. What are the important variables that influence the decision of how much of a good a consumer might wish to purchase over a given time period? Observation, introspection, and past investigations suggest that three variables are important: the price of the good in question, the price of related goods, and income. For the consumer, these variables are exogenous, that is, variables whose values are assumed by the consumer to be unaffected by his or her decisions. The endogenous variable in this sub-model is the quantity per unit time that is bought.

Logical structuring

Functional relationship We can express the logical relationship between these variables in a useful shorthand method by denoting each variable by a symbol and using functional notation to express the sentence: "The quantity demanded per unit of time by a consumer is influenced by the price of the good, by the price of related goods, and by the consumer's income". To be concrete, suppose the good in question is wine. Denote the quantity demanded per unit of time by Q_V^d, (the V is mnemonic for "vino"), the own price by P_V, the prices of related goods by, for example, P_B for the price of beer and P_F for the price of food, and income per unit of time by Y. The equation can be written as:

$$Q_V^d = f(P_V, P_B, P_F, Y) \tag{3.1}$$

and is to be read as: "The quantity of wine demanded per unit of time is a function of the price of wine, the price of beer, the price of food, and income per unit of time." The variables P_V, P_B, P_F, and Y are called the arguments of the function f. The statement that "the quantity demanded is a *function* of the variables P_V, P_B, P_F, and Y means that once particular numerical values are specified for each of these four variables then there is a unique associated numerical value for the variable Q_V. We will shortly elaborate on this concept.

A demand function is also referred to as a demand schedule. This terminology emphasizes that the relationship between the exogenous variables and the quantity demanded can also be expressed in the following way: A demand schedule tells how much of a good (per unit of time) would be bought at any possible values for prices and income. That is, for every set of values of exogenous variables, it "reads off", much like reading information from a schedule, the quantity that would be demanded.

An example of such a schedule might look like Table 3.1. In that table, column one (1) specifies values for the endogenous variable Q_V^d, and the other four columns specify values for the exogenous variables $P_v,\ P_B,\ P_F$,and Y The values for the three exogenous variables $P_B,\ P_F$,and Y are set at one (1), two (2), and one (1), respectively. Three different values of the exogenous variable P_V are stipulated: zero (0),one (1), and two (2). For each set of values for the exogenous variables, e.g., $\{P_V = 1, P_B = 1, P_F = 2, Y = 1\}$, the quantity demanded per unit of time associated with that set is specified in the first column.

Q_V^d	P_V	P_B	P_F	Y
2	0	1	2	1
1	1	1	2	1
0	2	1	2	1

Table 3.1

Now, the information captured by our notion of the demand relationship could be captured in a collection of tables like Table 3.1, each table having a different set of values for exogenous variables and associated values for the endogenous variable. There are two problems with expressing the information embodied in our equational statement of demand via such a set of tables. First, the tables are too restrictive for our purposes. Our theory, for reasons that will be clearer after the next chapter, is not usually capable of predicting exact relationships between actual numbers. Rather, it is a qualitative prediction. Hence, a table of values can only serve as an example, and cannot convey the general proposition about the variables embodied in our equational statement.

Second, even as an example, a table of values is cumbersome. First of all, for a market such as the one we are describing, the units for quantities per unit of time are surely on the order of millions of bottles per year. That is, when we stipulate "$Q_V^d = 2$" we surely should interpret the number 2 as measuring millions of bottles per year. Consequently, a table of values should have many more entries, ranging from one bottle to two million bottles. Second, even ignoring this problem, we would need a different table for each different set of values for the exogenous variables. Even with the values of the five variables restricted to whole numbers between zero and five, this becomes a massive number of tables.

The above table, though, can help us understand both the idea of a function and how we use the tools of analytic geometry to depict functions. Note that Table 3.1 keeps the values of the last three exogenous variables constant over all four rows. The only variation is between the values of the exogenous (to the consumers) own price P_V and the quantity demanded per unit of time, Q_V^d. That is, when we look at only the two variables whose values change row to row, we have *a set of ordered pairs of numbers* (P_V, Q_V^d) such that to each value of the first variable there corresponds a unique value of the second variable. The set of ordered pairs from our example is $\{(0,2), (1,1), (2,0)\}$. A *set* can also be thought of as a collection, and standard notation uses the curly braces $\{\}$ to identify a set, and uses parenthetical brackets () to denote members of a set. The idea of *ordered* pair is that the first element of any pair always represents the same variable, and the second element also always represents the same (albeit different from the first variable) variable.

The formal mathematical definition of a function of two variables is in fact just this: a set of ordered pairs of numbers (P_V, Q_V^d) such that to each value of the first variable P_V there corresponds a unique value of the second variable Q_V^d. We emphasize that a two-variable function is a set of ordered pairs of numbers because analytic geometry gives us tools to depict ordered pairs of numbers, namely points (and collections of points known as curves) in the Cartesian plane.

Graphical representation Analytic geometry allows us to picture ordered pairs of numbers and algebraic equations in terms of points and geometric curves. For many people, this depiction is the key tool for understanding logical interactions among variables. The key idea, the discovery of which is credited to the French mathematician Descartes (1596-1650) involves locating a point in a plane by means of its distance from two perpendicular axis. Such a plane is known as a "Cartesian plane" and points in this plane are located by pairs of numbers known as "Cartesian coordinates." This terminology is in commemoration of Descartes. We briefly review these concepts before developing graphical representations of the logical relationships that make up our sub-model of demand.

Coordinates The fundamental idea in analytic geometry is the establishment of a one-to-one correspondence between numbers or groups of numbers and points in a geometric space. "One-to-one" means that for every unique point there corresponds a unique pair of numbers. Of most use to undergraduate-level econmics is the correspondence between points in a plane and pairs of numbers. For our partial-equilibrium demand-supply model, the pairs of numbers of interest are (P_V, Q_V). Generic notation familiar to some from mathematics classes denotes pairs as (x, y). Most people are familiar with this basic concept from knowledge of map coordinates. A point on a map is described by it's *coordinates* : a pair of numbers, one of which specifies latitude and the other longitude.

To establish the one-to-one correspondence between points in a plane and pairs of numbers, start with a horizontal line in a plane, extending indefinitely to the left and the right. In generic notation, this line is known as the $x-axis$. For the model of this section, we might want to think of this axis as the P_V-axis (although later in this section, for quirky reasons of historical developments in economic thought, we will "switch" the notation of this axis). A reference point O on this axis and a unit of length (e.g., price of a bottle of wine) are then chosen. The axis is scaled by this unit of length so that the number zero is attached to point O, the number $+a$ is attached to the point a units of length to the right of O, the number $+2a$ is attached to the point $2a$ units to the right of O, the number $-a$ is attached to the point a units to the left of O, and so on. In this way, every point on the $x - axis$ corresponds to a unique real number. A real number is any number that can be expressed as a non-repeating or repeating decimal number, e.g., 1.625000... or 2.1417... .The perhaps non-intuitive feature of this real number line is that between any two real numbers, no matter how close to each other, there can always to interspersed another real number. This implies that a point on the line takes up no space. For the purposes of economic models, one can think of the real number line as a convenient approximation to numbers that represent small but discrete units.

Now place another straight line vertical in the plane, i.e., at a right angle to the x-axis, through point O. In generic notation, this is the $y-axis$. For the model of this section, we (initially) would think of this as the Q_V-axis. It also extends indefinitely up and down. Choose a unit of length, such as quantity of wine per unit of time, and scale the y-axis with this unit, much as with the x-axis. That is, the number b is attached to the point on the y-axis b units above point O, the number $2b$ is attached to the point on the y-axis $2b$ units above point O, and so on.

Now draw a line parallel to the $y-axis$ through point a on the $x-axis$ and a line parallel to the $x-axis$ through point b on the $y-axis$. These two lines intersect at point R, which corresponds to the pair of numbers (a,b). Clearly, for any two real numbers a and b, there corresponds a unique point R, which we denote as $R(a,b)$. Conversely, we say that the coordinates of R are (a,b). Figure 3.2 illustrates the point associated with the pair $(2,1)$.

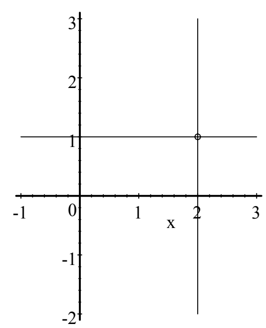

Figure 3.2: The Cartesian Plane

The two axis divide the plane into four quadrants labeled I, II, III, and IV, where standard terminology denotes quadrant I as that section for which all points have two positive coordinates. This quadrant is of most use in economics because most economic variables, such as prices and quantities, are inherently non-negative numbers. Quadrant II has points with signs $(-,+)$, quadrant III has points with signs $(-,-)$ and quadrant IV has points with signs $(+,-)$.

Graphs of functions With the establishment of a one-to-one relationship between pairs of numbers and points in a plane, we can now depict the functional relationship between Q_V^d and P_V captured in Table 3.1 (for the stipulated fixed values of the other exogenous variables) and expressed as the set of ordered pairs $\{(0,2),(1,1),(2,0)\}$. The first member of the pair is measured along the horizontal axis and the second member along the vertical axis. These points are depicted in Figure 3.3.

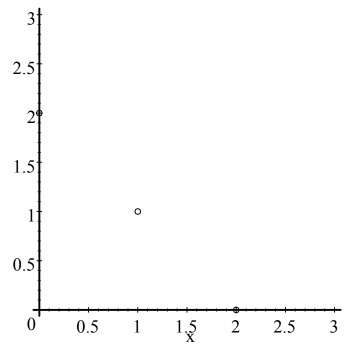

Figure 3.3: Graph of Table 3.1.

As noted, expressing this information in table form is cumbersome and limiting. Consequently, we like to express this information in equational form, either in words or in functional notation so as to succinctly capture the relationship between the endogenous and exogenous variables in the demand relationship in a symbolic form. We also like to allow the variables to take any value in some specified segment of the real number line, because this facilitates representation of relationships as continuous curves. This representation provides a much more concise way of expressing all the members of a function which has many members, as do most of the functions in which we are interested. A useful example of what we are talking about here is a *linear* function. To illustrate the idea of a linear function, note that the tabular example we have been working with has three points, all of which could be members of a function described as follows:

Definition 1 *Let P_V be allowed to take on any number drawn from the set of real numbers that are between and include zero (0) and two (2). The function $Q_V^d = f(P_V)$ is the set of all ordered pairs (P_v, Q_V^d) such that $0 \leq P_V \leq 2$ and*

$$Q_V^d = 2 - P_V$$

We can immediately check that the three points from our table are members of this function by substituting the three values of P_V into the equation $Q_V^d = 2 - P_V$ and confirming that the values of Q_V^d that result do indeed correspond to the associated points from the table. From this function, we could construct as detailed a table as we would like by picking values of P_V and subtracting this value from two (2) to get the associated value of Q_V^d. If we wanted to construct a table that incorporated every possible pair, though, it would be an infinitely large table, because there are an infinite number of real numbers that lie in the interval between zero and two. This means that the functional relationship will have an infinite number of set members. The graph of this set of points, though, is easily depicted in the Cartesian plane as a continuous straight line between the points $(0, 2)$ and $(2, 0)$. We can think of this graph as if constructed by putting a straight-edge on graph paper between the two points $(0, 2)$ and $(2, 0)$ and drawing a line with a pencil along this edge without lifting the pencil from the paper. Such a graph is depicted in Figure 3.4.

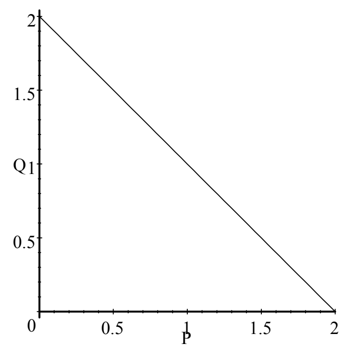

Figure 3.4: Demand Curve

As noted earlier, the point here is not that we think most economic variables are measured continuously. The assumption of continuous measurement is simply convenient and 'innocuous.' An innocuous assumption is one that can be replaced by a more realistic or complicated assumption without affecting the main conclusions or implications drawn from the model. When we describe an assumption as innocuous, we are appealing to the authority of the economics profession at large, members of which have in fact determined through research that the assumption doesn't affect the main conclusions of the theory.

In generic notation, we frequently write a linear function as follows:

$$y = mx + b$$

where y and x are variables and m and b are numbers known as the parameters of the function. When written in this fashion, we say the function is written in slope-intercept form. The y-axis intercept is the value of y when $x = 0$. The slope of the function is symbolized by m , and tells you the change in the y-variable for any given change in the x-variable as you move from one point on the line to another. To elaborate, suppose you know that a point $P_1(x_1, y_1)$ lies on the above line and a point $P_2(x_2, y_2)$ also lies on the line, where the x_i's and y_i's stand for particular numerical values, such as $x_1 = 1,\ x_2 = \frac{1}{2}, y_1 = 1,\ y_2 = 1\frac{1}{2}$.For these particular values, the two points P_1 and P_2 would be $(1, 1)$ and $(\frac{1}{2}, 1\frac{1}{2})$, respectively. The *change in* y from point P_2 to point P_1, symbolized as either $y_2 - y_1$ or Δy, is defined as the value of the $y-$coordinate of point P_2 minus the value of the $y-$coordinate of point P_1. The *change in* x, symbolized as either $x_2 - x_1$ or Δx, is defined as the value of the $x-$coordinate of point P_2 minus the value of the $x-$coordinate of point P_1. For the example points $P_1(1, 1)$ and $P_2(\frac{1}{2}, 1\frac{1}{2})$, $\Delta y = \ \frac{1}{2}$, and $\Delta x = -\frac{1}{2}$. For any arbitrary value of

$y \neq y_1$, and arbitrary value of $x \neq x_1$, the ratio $\frac{y-y_1}{x-x_1}$ is known as the slope, m. For the example of our linear demand curve, the intercept is 2 and the slope is -1. Note that the change in the y-value coordinates of any two points on the line divide by the change in x-value coordinates for these two same points is always the same number, m.

With this background in how relationships between variables are represented and depicted, we move on to make additional behavioral assumptions about demand. We also use this as an opportunity to generalize the idea of a function from a relationship between just two variables to a relationship among many variables

Logical structuring: additional behavioral assumptions Now that the decision has been made about what variables are important and which of these variables are endogenous and exogenous, more specific behavioral assumptions can be made. More than just specifying what variables affect the quantity demanded, we sometimes can specify in what direction these variables affect the quantity demanded: for example, we might be able to answer the question, if P_V goes up, how does Q_V^d change?

Let us go through the list of exogenous variables and recount what the theory learned in principles courses tells us about the qualitative impact of hypothetical changes in the value of these variables on the quantity demanded per unit of time. For each variable, the thought experiment being conducted holds constant the values of the other exogenous variables. This is called the *ceterus paribus* assumption: ceterus paribus is a Latin phrase meaning "other things being equal".

1. Own price. Ceterus paribus, as a good's own price goes down, the quantity demanded per unit of time is expected to increase. This inverse relationship between own price and quantity demanded is know as the Law of Demand. The example from the preceding section was chosen to reflect this assumption.

2. Prices of related goods. Ceterus paribus, as the price of a related good goes down, the quantity demanded per unit of time of the good in question goes up or down depending on whether the related good is a substitute or complement. For example, if P_B were to go down, one might expect the quantity demanded of wine to decrease as consumers' substitute purchases of the now relatively cheaper beer for wine. If this were the case, wine and beer would be considered substitutes. On the other hand, if P_F were to go down, one might expect the quantity demanded of wine to increase; since people do tend to consume wine and food together, a lower price of food, via the Law of Demand, would lead to higher food consumption and consequently higher wine consumption.

3. Income. Ceterus paribus, as income goes down, the quantity demanded per unit of time goes up or down depending on whether the good in question is a normal good or an inferior good. For a normal goods, a reduction in income induces people to reduce consumption of the good; for an inferior good, an income reduction increases consumption of the good.

We can succinctly display the above information about behavior by putting a "plus" or a "minus" above each argument in the demand function to denote the direction of the change in quantity demanded per unit of time induced by an increase in the argument, ceterus paribus. For example, if wine is a normal good, and beer is a substitute for wine while food is a complement, we would write the quantity demanded relationship as

$$Q_V^d = f(\overset{-}{P_V}, \overset{+}{P_B}, \overset{-}{P_F}, \overset{+}{Y}) \tag{3.2}$$

Those familiar with calculus will recognize that the above information is also captured by the stipulated algebraic signs of the partial derivatives of the function f (assuming f has those properties that make differentiation acceptable and possible):

$$\frac{\partial Q_V^d}{\partial P_V} < 0; \ \frac{\partial Q_V^d}{\partial P_B} > 0; \ \frac{\partial Q_V^d}{\partial P_F} < 0; \ \frac{\partial Q_V^d}{\partial Y} > 0. \tag{3.2'}$$

We can also graphically represent this more completely specified model of demand. The problem is that we have five variables, but only two axis to work with in a two-dimensional picture. In our previous example built around a hypothetical tabular representation of a demand schedule, we assumed the exogenous variables P_B, P_F, and Y were given constant values of 1, 2, and 1, respectively, and only values for P_V and Q_V^d varied. Therefore the graph of the relationship between the own-price P_V and quantity Q_V^d simply depicted the two columns with values for those variables. We could have, though, looked at a table with constant values for the exogenous variables P_V, P_F, and Y. and varying values of Q_V^d and P_B. Such a table might look as follows:

Q_V^d	P_V	P_B	P_F	Y
2	1	2	2	1
1	1	1	2	1
0	1	0	2	1

Table 3.2

Alternatively, we could have held constant P_V, P_B, and P_F and let Q_V^d and Y vary. Or we could have held constant P_V, P_B, and Y and let Q_V^d and P_F vary. Hence, in a two-dimensional graph, we could depict the relationship between the endogenous variable Q_V^d and any one of the the other exogenous variables, much as we did with Q_V^d and P_V

The reason we plot P_V instead of any of the other exogenous variables on the vertical axis is because P_V is, for this model, *conditionally* exogenous. That is, it is exogenous to this sub-model, but not to the complete model. The other exogenous variables are not only exogenous to the sub-model, but are exogenous to the model as a whole. Choosing to depict the demand relationship in terms of P_V and Q_V^d on the two axis of a two-dimensional diagram reflects the modeler's knowledge that P_V is ultimately a variable whose value we wish to determine within the model.

We point this out to emphasize that the choice of how a relation is depicted graphically is often made for pedagogical reasons that can only be known on the basis of knowledge of how the sub-model fits into the larger model. To someone learning the model for the first time, it is unlikely that the reasons for a particular depiction will be obvious. What is important for the student at this stage of model development is to understand the ultimate goals of the model- what variables are to be explained- and the logical correctness of the intermediate steps involved in presenting the sub-model.

We also should point out that we focus on thinking about pairs of variables that are allowed to vary while the values of other exogenous variables are held constant simply because we want to develop our model in a 2-dimensional-friendly format. Our equation (3.1) expresses Q_V^d as a function of four exogenous variables; we can only turn it into a two-variable function by the artifice of imagining that we are holding constant the values of three of these four exogenous variables.

The basic idea behind a two-variable function is that once we know the value of an exogenous variable we then know the value of the endogenous variable. This basic idea extends to the multi-variable case as well. In the example we are working with, the notion that Q_V^d is a function of P_V, P_B, P_F and Y means that once numbers are known for the values of these four exogenous variables, then the value of Q_V^d is known. These five numbers thus form an ordered five-tuple $(P_V, P_B, P_F, Y, Q_V^d)$. The set of all such ordered five-tuples formed by substituting permissible values of the four exogenous variables and the corresponding value of Q_V^d is a function. Thus, much like in the two-variable case, we see that a function is a set of points. The points here, though, are not two-dimensional but rather five-dimensional. This is hard to visualize.

To gain some insight into why we refer to the ordered five-tuple as a point, think about the visual representation of a three-variable function that most of us have had some experience with: a topographical relief map. Such a map is constructed by forming a two-dimensional grid along whose axis are measured latitude and longitude. Every point on this flat grid thus is represented by an ordered pair whose elements are latitude and longitude coordinates, respectively. Now, superimposed on top of this grid is a physical replica of the topography of the area-mountains, valleys, and the like. If the map is a good one, for each map coordinate, the physical topographical map has the same elevation (to scale) as does the real physical feature. This means that for every ordered pair that plots a point in terms of latitude and longitude, we can associate a third number, the elevation of the terrain at that point. The surface of the relief map is just a depiction of the ordered triplets that make up the function that relates terrain elevation to latitude and longitude.

Even though we can visualize and even draw on a two-dimensional surface a picture of a three-variable function such as a relief map, it is still difficult. In fact, most maps are not drawn in three dimensions, but use devices such as "iso-elevation" lines to collapse into two dimensions the same information depicted in a three-dimensional graph. An iso-elevation line is the set of all those ordered pairs of latitude and longitude coordinates for which the associated elevation is the same. For example, imagine we have a relief map of part of a mountain. If, looking directly down from above onto the model of the mountain, we imagined someone (lliliputian size, of course) walking around the mountain by maintaining his or her elevation at a constant value, we would see a curve being traced out on the latitude-longitude plane. This curve would be what we call an iso-elevation curve, the "iso" prefix coming from the Greek "isos" which means "equals." Most people can envision the mountain just from looking at a family of iso-elevation lines. when the lines are close together, the mountain is relatively steep, and when they are far apart, the mountain is flatter. Figure 3.5 is a three-dimensional graph of the function $z = \sqrt{xy}$ that roughly corresponds to a "mountain." For this function to represent a mountain, z must denote elevation and x and y must denote latitude and longitude. Figure 3.6 depicts the information via "iso-elevation." lines, also known as contour lines. That is, each of the curves in Figure 3.6 connects those adjacent pairs (x, y) for which z has the same value.

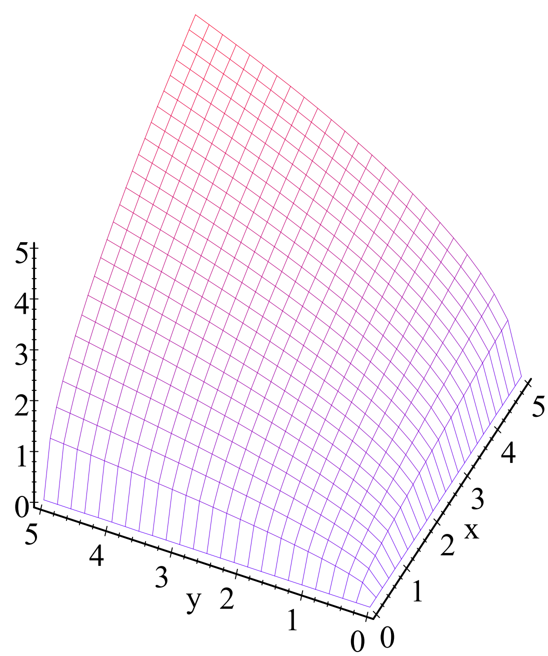

Figure 3.5.i: A Three-Variable Function Graph

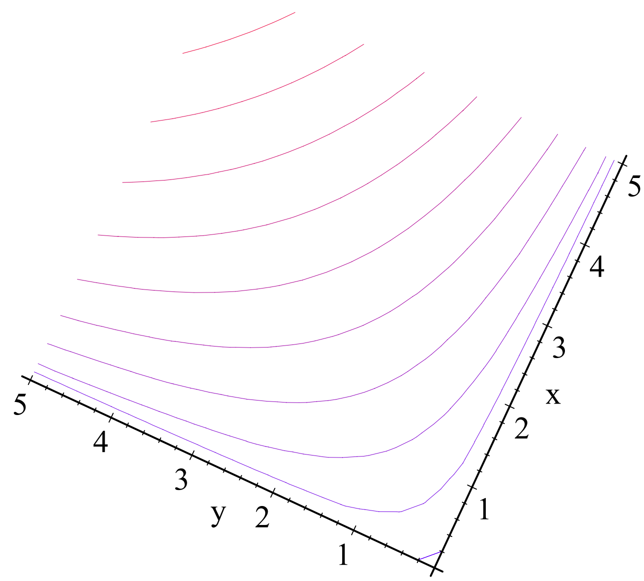

Figure 3.5.ii: Contours of a three-variable function

For this three-variable function, we can visualize the ordered triplets that are its members. They constitute the surface of the "mountain." The point we want to emphasize, though, is that the members of the set are ordered triplets. While we can't visually depict ordered four-tuples, five-tuples, and so on, we can still usefully think of them as "points" analogous to the three-variable and two-variable case.

Depicting the additional assumptions: curve-shifting. How, then, do we collapse the information in our multi-variable demand function into a two-dimensional graph? We do this in two ways. In the first approach, we impose the ceterus paribus assumption on the three exogenous variables P_B, P_F, and Y. We then plot in the two-dimensional $P_V - Q_V^d$ plane the curve relating Q_V^d to P_V for these given other exogenous variable values. Finally, we see how the placement of this curve in the $P_V - Q_V^d$ plane changes as we keep two of the three other exogenous variables constant and look at different values of just one of the three other exogenous variables. In the second approach, we make use of something similar to the "iso-elevation" lines used to collapse the information in a relief map onto a two-dimensional map.

Assumption of specific functional forms To gain insight into this procedure, it helps to illustrate the process with a few specific examples. To help make an abstract equation like (3.1) or (3.2) more concrete, we frequently specify the relationship as a particular equation. This is known as picking a specific functional form to serve as an example. We will work through two examples, each with a different but specific functional form.

The linear case First consider a linear relationship:

$$Q_V^d = a_0 - a_1 P_V + a_2 P_B - a_3 P_F + a_4 Y,\ a_i \geq 0,\ i = 1, 2, 3, 4. \tag{3.3}$$

The $a_i's$ in (3.3) are examples of parameters. As noted, they are like exogenous variables in that they symbolize numbers that are determined outside of the model, and they are usually assumed to be constant. All of them are restricted to be non-negative so that equation (3.3) conforms to the assumptions we made about how Q_V^d moves in response to changes in the values of the exogenous variables, i.e., as P_F gets larger, ceterus paribus, Q_V^d gets smaller.

At this point, we need to address one of the peculiarities of the economics profession in terms of how it translates equations into graphs. In keeping with a long tradition in economics, the price of wine, measured in units of currency/unit of wine, is measured along the vertical axis, while the quantity/unit of time is measured along the horizontal axis. Note that this means we are measuring the endogenous variable Q_V^d along the horizontal axis and the exogenous variable P_V along the vertical axis. This tradition is sometimes confusing to students who have been schooled in high school algebra to graph functions with the dependent (endogenous) variable on the vertical axis and the independent (exogenous) variable on the horizontal axis, as we did in Figures 3.3 and 3.4.

For the linear functional form of (3.3), one can use algebra to express the demand relationship with P_V on the left-hand-side of the equality sign and all other variables on the right:

$$P_V = (\frac{a_0}{a_1}) - (\frac{1}{a_1})Q_V^d + (\frac{a_2}{a_1})P_B - (\frac{a_3}{a_1})P_F + (\frac{a_4}{a_1})Y. \tag{3.4}$$

We call this the *inverse form* of the demand curve. In general functional notation, we would write this as follows:

$$P_V = f^{-1}(\overset{-}{Q_V^d}, \overset{+}{P_B}, \overset{-}{P_F}, \overset{+}{Y}) \tag{3.5}$$

where again the algebraic signs above each argument tells us the direction in which the value of left-hand- variables moves when the value of a right-hand-side argument goes up. These signs can be inferred from our fundamental assumptions embodied in equation (3.2). For the linear functional form, this is straightforward: we can infer the signs by looking at the signs of the coefficients on the right-hand-side of (3.4). The symbol f^{-1} is the traditional way of describing an inverse function, and is just another symbol that is to be read "...is a function of...". Again, let us emphasize that this inverse form better corresponds to the traditional graphical approach in economics of measuring own price on the vertical axis and quantity demanded per unit of time on the horizontal axis.

We could go one step further in making our examples concrete by specifying actual numbers as examples of parameters. For example, we could specify a linear demand curve as:

$$Q_V^d = .5 - P_V + .5P_B - (.5)P_F + 2Y. \tag{3.5}$$

Occasionally, this further step away from the abstract can provide a useful mental hook on which to hang a concept. It also helps us understand the pedagogical technique of "curve-shifting" that economists frequently use to teach models. Economists will frequently say something along the lines of : "If, ceterus paribus, income increases, this shifts the demand curve out to the right." To understand better what this means, first rearrange equation (3.5) in inverse form:

$$P_V = .5 - Q_V^d + .5P_B - (.5)P_F + 2Y. \tag{3.6}$$

Now imagine that hypothetical numerical values for the three exogenous variables the price of beer, the price of food, and the value of income are $P_B = 1$, $P_F = 2$ and $Y = 1$. The demand curve represented by equation (3.5) could now be written as

$$\begin{aligned} Q_V^d &= .5 - P_V + .5\overbrace{(1)}^{P_B} - (.5)\overbrace{(2)}^{P_F} + 2\overbrace{(1)}^{Y} \\ &= 2 - P_V. \end{aligned} \tag{3.7}$$

Equivalently, (3.6), would be

$$P_V = 2 - Q_V^d. \tag{3.8}$$

With these numerical values for the exogenous variables, the inverse form is a straight line with slope of minus one (-1) and intercept of two (2). This is depicted in Figure 3.6.

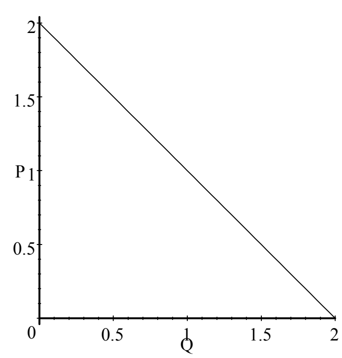

Figure 3.6: Inverse Demand Curve

Now imagine a new hypothetical value of income of $Y = 1.25$,but with all other exogenous variables at their previously stipulated values. The graphical representation of this new curve is again a straight line with slope minus one (-1), but with intercept 2.5. This new curve is parallel to the old one, but has "shifted up" or, in equivalent language, "shifted out" so that associated with every value of Q_V^d there is now a higher value of P_V. These two curves are actually depicted in Figure 3.7, where the solid line depicts the case where $Y = 1$ and the red line depicts the case where $Y = 1.5$.

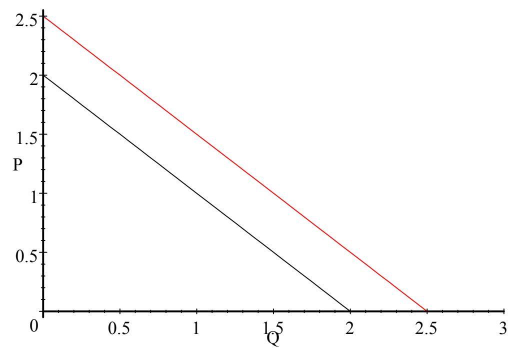

Figure 3.7: Two Inverse Demand Curves

An alternative way of thinking about how demand curves shift is analogous to thinking in terms of iso-elevation lines for the "mountain" graph. Contemplate holding constant the values of the exogenous variables P_B and P_F . To be less abstract, hold them constant at $P_B = 1$ and $P_F = 2$, as in the previous example. The demand curve (3.5) can now be written as

$$
\begin{aligned}
Q_V^d &= .5 - P_V + .5\overbrace{(1)}^{P_B} - (.5)\overbrace{(2)}^{P_F} + 2Y && \text{(3.9.i)} \\
&= -P_V + 2Y && \text{(3.9.ii)}
\end{aligned}
$$

We can now "solve out" for Y as a function of P_V and Q_V^d:

$$
Y = \left(\frac{1}{2}\right) Q_V^d + \left(\frac{1}{2}\right) P_V \tag{3.10}
$$

This is a three-variable function, and the picture of it is a three-dimensional plane, as illustrated by the colored surface in Figure 3.8.i.

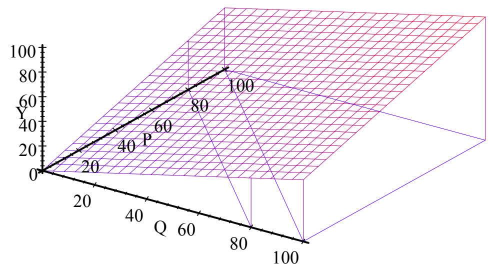

Figure 3.8.i: The $P_V - Q_V^d - Y$ plane

For some people, this is not an easy picture to visualize. The vertical axis measures Y and the two horizontal axis measure P_V and Q_V^d. You might want to think of this plane as a flat tilted awning spread over the first quadrant of the two-dimensional $Q_V^d - P_V$ plane. The dashed lines in the colored plane represent points in the three-dimensional plane for which Y is a constant value. The vertical "posts" along the side edges that appear to support the "awning" help one visualize how the height of the awning increases as it moves away from its base at $(Q_V^d = 0, P_v = 0, Y = 0)$: The height of the post is the value of Y at that point in the $Q_V^d - P_V$ plane. For example, at both $(Q_V^d = 0, P_v = 80)$ and $(Q_V^d = 80, P_v = 0)$, $Y = 40$, while at both $(Q_V^d = 0, P_v = 100)$ and $(Q_V^d = 100, P_v = 0)$, $Y = 50$. Finally, the highest "post" is situated at $(Q_V^d = 100, P_v = 100)$, and has the associated value of $Y = 100$.

Within this plane, there are 'contour lines' that represent three-dimensional points, i.e., ordered triplets (Q_V^d, P_V, Y) for which Y is constant. Figure 3.8.ii displays the plane with these contour lines shown against a solid background that represents the whole plane. For example, the line in the plane between the points $(80, 0, 40)$ and $(0, 80, 40)$ is the set of all those points for which P_V and Q_V^d are non-negative and for which $Y = 40$. We call these lines within the plane that have a constant value of Y *iso-income lines*.

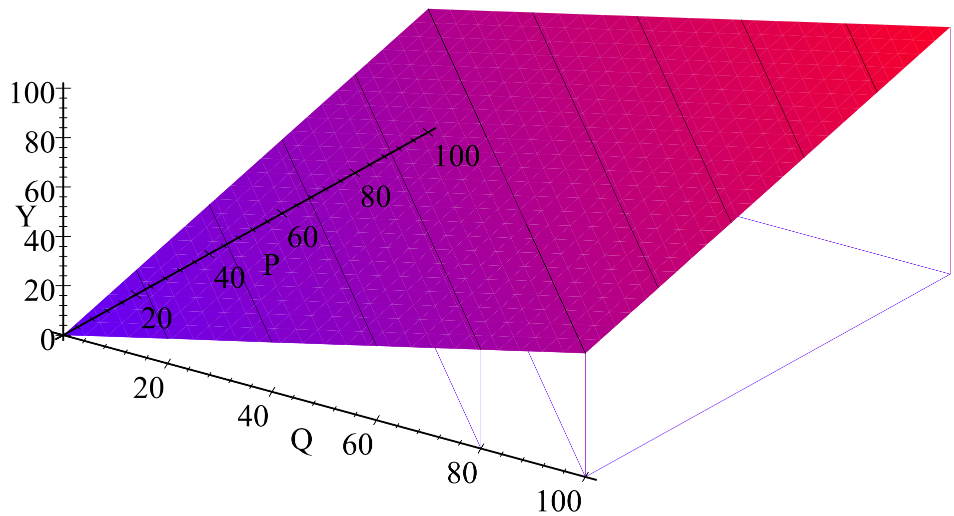

Figure 3.8.ii: Level contours of the $P_V - Q_V^d - Y$ plane

Now imagine what one would see if one peered down from above the "awning" onto the $P_V - Q_V^d$ plane. Much as in the case of a two-dimensional map with iso-elevation lines, one would see straight lines connecting the pairs of points in the $P_V - Q_V^d$ plane for which Y is some constant value. These are depicted in figure 3.8.iii.

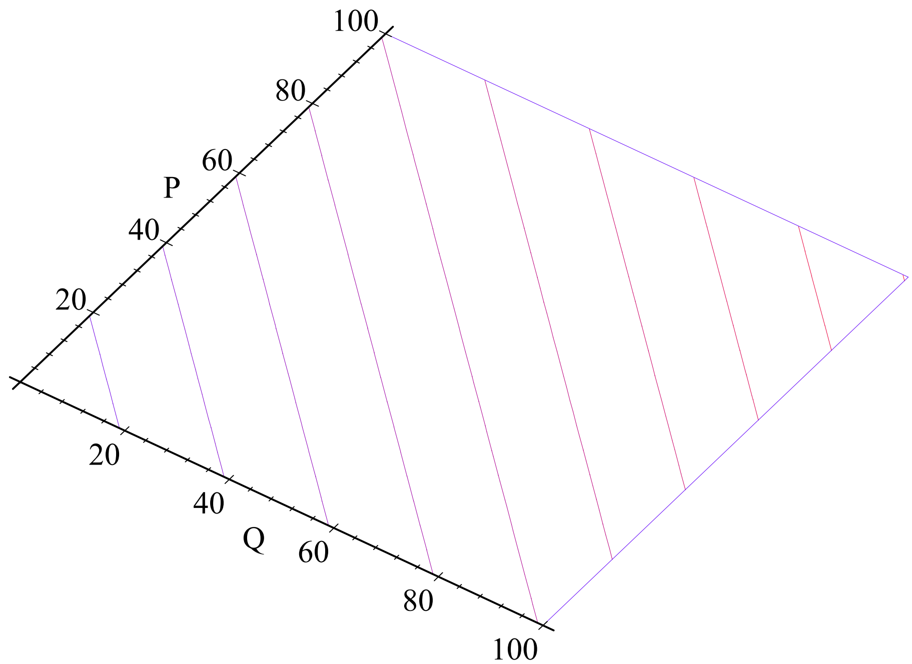

Figure 3.8.iii: Level lines of the $P_V - Q_V^d - Y$ plane

Each of these straight lines in the $Q_V^d - P_V$ plane represents a linear relationship between P_V and Q_V^d. For example, the line that connects the points $(80, 0)$ and $(0, 80)$ is

$$P_V = 80 - Q_V^d$$

and the line that connects $(1, 0)$ and $(0, 1)$ is

$$P_V = 2 - Q_V^d.$$

From this perspective, "shifts" in the demand curve (or, as here, the inverse demand curve) as drawn in the two-dimensional $Q_V^d - P_V$ plane can be thought of as movements from one "iso-income" line to another. We can think of the $Q_V^d - P_V$ plane as populated by a family of inverse demand curves, where each family member is associated with a particular value of Y.

Similar shifts of the demand curve arise from hypothetical changes in P_B holding P_F and Y constant, or from changes in P_F, holding P_B and Y constant. Whether the curve shifts out (right) or in (left) along the horizontal axis depends on whether the goods are substitutes or complements. For the case of beer, a substitute, an increase in the price of beer makes people substitute wine for beer, ceterus paribus. This means that for an unchanged level of income, an unchanged price of food, and an unchanged price of wine, people will consume more wine. This means the demand curve shifts out to the right.

Perhaps working out this example of a ceterus paribus change in P_B will drive home this important idea of curve-shifting. Now again imagine that we start with the same hypothetical numerical values for the three exogenous variables that we used in the "change in income" example: $P_B = 1$, $P_F = 2$ and $Y = 1$. The demand curve represented by equation (3.5) gain would be identical to (3.7), repeated here for convenience:

$$\begin{aligned} Q_V^d &= .5 - P_V + .5\overbrace{(1)}^{P_B} - (.5)\overbrace{(2)}^{P_F} + 2\overbrace{(1)}^{Y} \\ &= 2 - P_V. \end{aligned} \tag{3.7}$$

Equivalently, (3.6), would be

$$P_V = 2 - Q_V^d. \tag{3.8}$$

Again, with these numerical values for the exogenous variables, the inverse form is a straight line with slope of minus one (-1) and intercept of two (2).

Now suppose P_F and Y keep the same values, i.e., $P_F = 2$ and $Y = 1$, but now $P_B = 2$. Substituting these values into (3.7) gives us the new equation

$$Q_V^d = .5 - P_V + .5\overbrace{(2)}^{P_B} - (.5)\overbrace{(2)}^{P_F} + 2\overbrace{(1)}^{Y} \tag{3.11.i}$$

$$= 2.5 - P_V. \tag{3.11.ii}$$

Note how the value of the intercept has changed from 2 to 2.5, but how the slope has not changed. The curve has shifted out, or equivalently, shifted up, just as in the just-discussed case of an increase in Y. The fact that only the intercept and not the slope has changed is not a general feature of these kinds of exercises. What is general is that, for every permissible value of P_V , the associated value of Q_V^d is now bigger than it was.The multiplicative case

To make sure this concept of curve-shifting is clear, now consider another example in which we use a different functional form. Consider the following specification:

$$Q_V^d = \alpha_0 (P_V)^{(-\alpha_1)} (P_B)^{(\alpha_2)} (P_F)^{(-\alpha_3)} (Y)^{(\alpha_4)}, \alpha_i \geq 0, i = 0, 1, 2, 3, 4. \tag{3.12}$$

where the $\alpha_i' s$ are parameters. Remembering the rules of exponents from high school algebra, this can be expressed in inverse form as

$$P_V = (\frac{Q_V^d}{\alpha_0})^{(\frac{-1}{\alpha_1})} (P_B)^{(\frac{\alpha_2}{\alpha_1})} (P_F)^{(\frac{-\alpha_3}{\alpha_1})} (Y)^{(\frac{\alpha_4}{\alpha_1})}. \tag{3.13}$$

We will see that this example of a demand curve, like the linear example above, satisfies the general assumptions we made about how the quantity demanded per unit of time is affected by the price and income variables. That is, the quantity demanded changes in the stipulated direction for an increase in each of the price and income variables.

As with the linear example, let us assume the following specific numerical values for the $\alpha_i's$ and the exogenous variables other than P_V : $\alpha_0 = 1,\ \alpha_1 = 1,\ \alpha_2 = 2,\ \alpha_3 = 2,\ \alpha_4 = 1;\ P_B = 1,\ P_F = 2,\ Y = 1.$ With these values, equation (3.13) becomes:

$$P_V = \Big(\underbrace{\frac{Q_V^d}{1}}_{\alpha_0}\Big)^{\overbrace{(-1)}^{\frac{-1}{\alpha_1}}} \overbrace{(1)}^{P_B}{}^{\overbrace{(2)}^{\frac{\alpha_2}{\alpha_1}}} \overbrace{(2)}^{P_F}{}^{\overbrace{(-2)}^{\frac{-\alpha_3}{\alpha_1}}} \overbrace{(1)}^{Y}{}^{\overbrace{(1)}^{\frac{\alpha_4}{\alpha_1}}} \qquad (3.14.i)$$

$$= \frac{1}{4Q_V^d} \qquad (3.14.ii)$$

For someone less familiar with a function of this form, it may help to plot a few points to help envision the relationship. For example, let $Q_V^d = \frac{1}{4}$. The function tells us that the associated value is $P_V = 1$, so the point $(\frac{1}{4}, 1)$ is a member of the set that makes up this function. Now let $Q_V^d = 1$; the associated value is $P_V = \frac{1}{4}$, and the point $(1, \frac{1}{4})$ belongs to this function. Finally, consider $Q_V^d = 4$; the associated value is $P_V = \frac{1}{16}$, and the point $(4, \frac{1}{16})$ belongs to this function. Figure 3.9 displays the first two of these points along with plots of enough other points with values of Q_V^d ranging between 0.2 to 1.2 so as to present a picture of a smooth curve that represents this function.

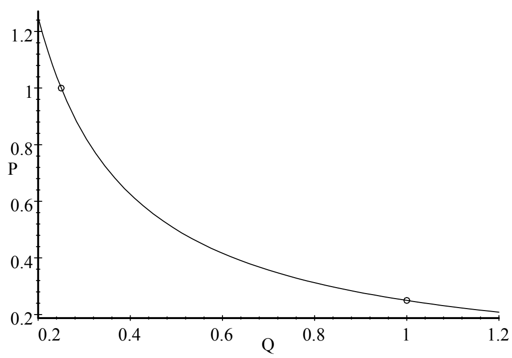

Figure3.9: Another demand curve example:

Now let us see how this curve shifts in response to a ceterus paribus change in one of the other exogenous variables such as Y. As with the linear example, now, instead of having $Y = 1$, let $Y = 2$, with all other exogenous variables remaining at their previous values. The inverse demand function then becomes

$$P_V = \frac{1}{2Q_V^d} \qquad (3.15)$$

Again, plot a few points for this function. For the values of Q_V^d that we just used when $Y = 1$, the new points that are members of this new function are

$(\frac{1}{4}, 2)$, $(1, \frac{1}{2})$ and $(4, \frac{1}{8})$. For each of these values of Q_V^d, the associated value of P_V is higher. This is true not just for these particular values of Q_V^d, but for any positive value of Q_V^d. Figure 3.10 superimposes the graph of equation 3.15 on the the graph of 3.14.ii to illustrate this "shift."

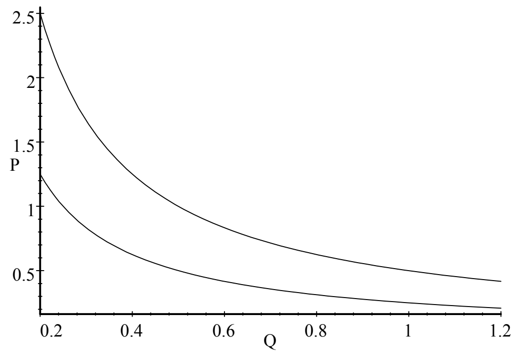

Figure 3.10: An upward shift

"Curve-shifting" in this multiplicative functional form case can also be thought of as moving from one "family member" to another. Note that we could write equation 3.14 with Y as a variable instead of assigning it the value of 1 (one). In this case the inverse demand curve would be

$$P_V = \frac{Y}{4Q_V^d} \tag{3.14.iii}$$

Multiplying both sides by $4Q_V^d$ lets us "solve out," for Y as a function of Q_V^d :

$$Y = 4P_V Q_V^d \tag{3.14.iv}$$

This 3-variable function is displayed in Figure 3.11.i.

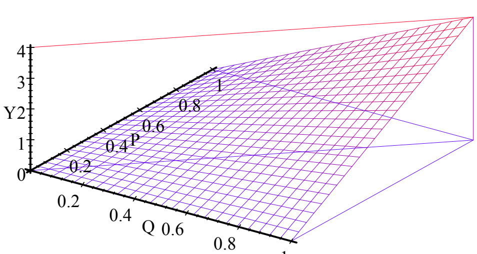

Figure 3.11.i: Surface of $Y = 4P_V Q_V^d$

As with the linear case, there are iso-income 'contour lines' that represent three-dimensional points, i.e., ordered triplets (Q_V^d, P_V, Y) for which Y is constant. Figure 3.11.ii displays the plane with these contour lines shown against a solid background that represents the whole surface.

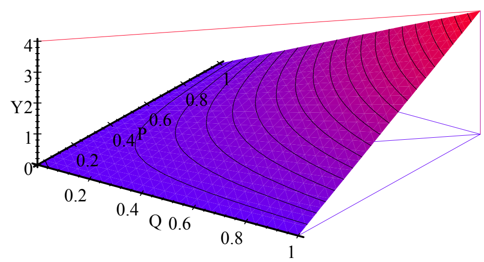

Figure 3.11.ii: Level contours of $Y = 4P_V Q_V^d$

Finally, just as with the linear case, we can now imagine peering directly down onto the $P_V - Q_V^d$ plane and seeing the plot of these iso-income lines. These are displayed in Figure 3.11.iii:

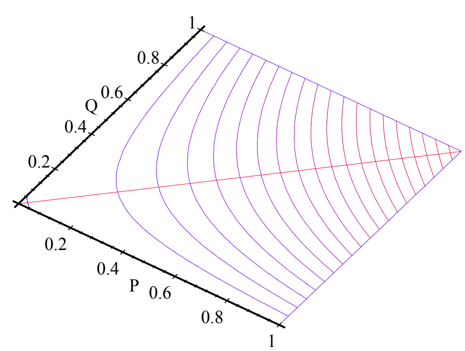

Figure 3.11.iii: Level curves of $Y = 4P_V Q_V^d$

Each of these curves is a two-variable function associated with a different value of Y. Just as with the linear case, then, we can think about different values of income, ceterus paribus, as creating different members of a family of demand curves.

Before moving on, we need to emphasize that the use of specific functional forms is mostly a pedagogical device: it permits us to make concrete the less constraining but harder-to-visualize general model. Specific functional forms are seldom an implication of theory. However, to be useful, they must be consistent with the theory. In the above demand curve examples, consistency of the specific functional forms with the theory is embodied in the sign restrictions on the a_i's and α_i's. Even though the different functional form assumptions gave rise to different graphs - one is a straight line, the other has curvature - both satisfy the theoretical restriction that requires the curve to slope down.

Also before leaving this introductory discussion of functional form, we note two other reasons besides the pedagogical one that economic models are expressed with specific functions. First, most general functions of interest to economists can be approximated by one or another specific functional form. Second, when economists use data to estimate an economic model, perhaps for forecasting purposes, they often need to assume a particular functional form. In fact, the "multiplicative case" of a demand curve is often labeled a "constant elasticity" demand curve because this functional form has a particular property- "constant elasticity"- that frequently fits economic data quite well. The hope of the economist is that the specific functional form chosen is in fact a good approximation to whatever the unknown "true" functional form might be.

Curve-shifting: the general case and "bi-directional"logic. As noted, often all that we know about a function are its qualitative properties, as described, for example, by equation (3.2). With only this limited qualitative information, how do we carry out the curve-shifting exercise that we just did with specific functional forms?

First note that by holding constant the values of the exogenous variables P_B, P_F, and Y at some particular values denoted by $P_{B,0}$, $P_{F,0}$, and Y_0, we have created a two-variable function that conceptually tells us an associated value of Q_V^d for every permissible value of P_V. The use of a numerical subscript, in this case zero or "naught," is standard notation for indicating that we are contemplating a particular value of a variable. We say "conceptually" to emphasis that we have no specific function in mind. We do know, though, that our behavioral assumption (1) implies that this function must associate lower values of Q_V^d with higher values of P_V . This means the graph of this function, with Q_V^d plotted on the vertical axis and P_V plotted on the horizontal axis, is a curve with negative slope. It also implies that the graph of the inverse demand function, with P_V measured on the vertical axis and Q_V^d measured on the horizontal axis, will also be a curve with negative slope.

In the tradition of economics, let us focus on the graph of the inverse demand curve when the values of the exogenous variables P_B, P_F, and Y are $(P_B)_0$, $(P_F)_0$, and Y_0, respectively. Contemplate now a particular value of Q_V^d, denoted by $(Q_V^d)_0$, and the associated value of P_V, denoted by $(P_V)_0$. Ask the hypothetical question: What would happen to Q_V^d if, at $P_V = (P_V)_0$, $P_B = (P_B)_0$, and $P_F = (P_F)_0$, Y were to increase, i.e., Y were to take a value greater than Y_0? Note that this hypothetical question simply imposes the ceterus paribus assumption on all the exogenous variables other than Y. The answer to our hypothetical question is thus given by our behavioral assumption that, ceterus paribus, an increase in income increases the quantity demanded. At the higher value of Y, it must be that $Q_V^d > (Q_V^d)_0$.

Now, there is nothing special about the value $(P_V)_0$. We could have carried out the same thought experiment for any of the permissible values of P_V, and for each value the same logic would apply: at the higher value of Y, it must be that $Q_V^d > (Q_V^d)_0$. Hence, at every value of P_V, the graph of the inverse demand function would have "shifted up." in response to an increase in Y.

We could also imagine going through the same exercise with the following change: instead of asking about what happens to Q_V^d at particular values of P_V when, ceterus paribus, Y increases, ask what must happen to P_V at particular values of Q_V^d when, ceterus paribus, Y increases This question is perhaps less intuitive because our behavioral assumption is that people *choose* quantity per unit of time based on the values taken by exogenous variables. This new question almost seems to ask: what price do people choose at a given value of Q_V^d when there is an increase in Y, ceterus paribus? But this is not correct. What we are really asking is: if we observe people consuming the same quantity in the face of an increase in Y (all other exogenous variables except P_V also being held constant, then what must have happened to P_V? Because people would consume more with higher Y and unchanged P_V, but we observe (or "hold constant" in our thought experiment) an unchanged Q_V^d, then it must be that P_V has gone up so as to offset the increased demand from the increase in Y. Hence, for every value of Q_V^d, the graph will have "shifted out."

To recap, with only information about the qualitative responses of people to changes in exogenous variables, we can still determine via thought experiments in which direction the demand curve will shift. We call this "bi-directional logic" because our strategy was to imagine holding constant either the variable measured on the vertical axis and asking what must happen to the variable measured on the horizontal axis, or holding constant the variable on the horizontal axis and asking what must happen to the variable on the vertical axis. In the former case, the curve "shifted out" while in the latter the curve "shifted up." Both depict the same phenomena.

Characterizing different tastes To this point, we have said nothing about "tastes". In our later more complete development of a complete model of consumer behavior, we will flesh out this concept in much more detail. Within the confines of this model, though, the concept of a consumer's tastes is embodied in the placement of the demand curve in the quantity-price plane. If, for example, we considered two consumers, Andy and Bob, who had identical incomes, and observed that Andy's inverse demand curve for wine was placed farther to the right than Bob's. This would say that at any hypothetical price of wine, given the price of related goods, Andy desires to purchase more wine per unit of time than does Bob. Since they both have the same income, this means that relative to Bob, Andy must prefer to consume more wine and less of some other good. Note that the concept of preference here necessarily implies a trade-off: because Andy "likes wine" more so than Bob, Andy must also *not like* some other good as much in comparison to Bob. To be clear about this, remember that, by assumption, Andy and Bob have identical incomes and are price-takers (the price of a good is assumed by Andy and by Bob to be unaffected by the quantity they purchase). Hence, for given prices of wine and other goods that apply to both Andy and Bob, if Andy buys more wine than does Bob, then Andy spends more of his income on wine than does Bob. Clearly, Andy must spend less than Bob on some other good, which in turn means he buys a smaller quantity. He likes this other good "less" than Bob. The partial-equilibrium model tends to obscure this observation.

The functional notation of equation (3.2) isn't very useful in helping one grasp the concept of how two consumers may have different tastes. For any individual, indexed by a subscript , we could denote that he or she has a unique demand curve by identifying each individual's function with a subscript:

$$Q^d_{V,l} = f_l(P_V, P_B, P_F, Y),\ l = 1, 2. \tag{3.16}$$

This still doesn't help one envision how the demand curves may differ. Again, to help make an abstract equation like (3.16) more concrete, we can specify the relationship as a particular equation. This, again, is known as picking a specific functional form to serve as an example. For example, we might again specify a linear relationship:

$$Q^d_{V,l} = a^l_0 - a^l P_V + a^l_2 P_B - a^l_3 P_F + a^l_4 Y,\ a^l_i \geq 0,\ i = 0, 1, 2, 3, 4; l = 1, 2. \tag{3.17}$$

The $a^{l\prime}_i s$ in (3.17) are examples of parameters, with the superscript l signifying a particular individual. As noted, they are like exogenous variables in that they symbolize numbers that are determined outside of the model, and they are usually assumed to be constant. They also make concrete what it means for different individuals to have different tastes: they would have different numerical values for some or all of the $a^{l\prime}_i s$.Supply

Now consider the sub-model for supply. What are the important variables that influence the decisions of firms about how much of a good to produce over a given time period? Twin underlying assumptions, one about behavior and the other about technology, motivate the answer to this question. One assumption describes the mainspring of firm motivation: firms attempt to maximize profits. A second assumption is that there exists a well-defined technology called the production function that tells how much output per unit of time can be produced for any given quantity of inputs, e.g., labor, per unit of time. Given these assumptions, a full-blown detailed model of firm decision-making tells us that the quantity produced per unit of time depends on the own price and the price of inputs into the production process. To be concrete, again assume we are concerned with production of wine. Denote the quantity of wine produced by Q^s_V. Assume the only inputs into the production process are labor and land, with associated prices w (for wage) and r (for rent). For the firm, the quantity produced is endogenous, and the price of wine and the price of inputs are exogenous. The shorthand expression for the relationship between these variables is the supply schedule

$$Q^s_V = g(P_V, w, r) \tag{3.18}$$

and can be read as "the quantity of wine produced is a function of the price of wine, the price of labor, and the rental price of land". As with the demand schedule, P_V, w, and r are the arguments of the function.

Now recall what principles courses (and a full-blown model of how the firm's decisions maximize profits) tell us about more specific behavioral assumptions concerning the relationships between these variables. As with the demand function, we can express these behavioral assumptions by placing the appropriate algebraic sign the relationships between these variables.

1. Own price. The higher the price of wine, the greater the quantity produced per unit of time, ceterus paribus.

2. Input prices. The higher the price of inputs such as labor, fuel, or raw materials, the smaller is the quantity produced per unit of time for any given price of wine.

As with the demand function, we can express these behavioral assumptions by placing the appropriate algebraic sign above each argument in the supply function:

$$Q_V^s = g(\overset{+}{P_V}, \overset{-}{w}, \overset{-}{r}). \tag{3.19}$$

The usual graphical depiction of this relationship is an upward-sloping curve with P_V on the vertical axis and quantity of wine per unit of time on the horizontal axis. Changes in w or r shift this curve in the price-quantity plane. The solid line in Figure 3.12 depicts such a schedule. The red line in Figure 3.12 depicts a leftward shift in the supply schedule that results from an increase in either w or r.

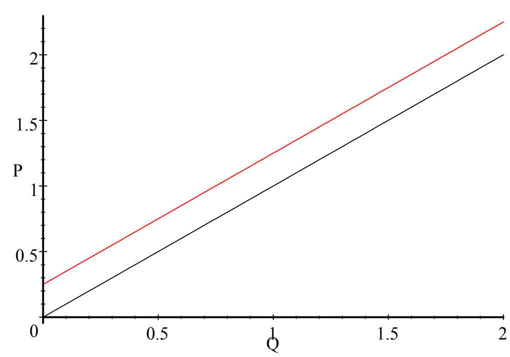

Figure 3.12: Supply Curve

Analogous to the discussion of tastes in the demand sub-model, production technology is captured by the specific placement of the supply curve in the price-quantity plane. As in that discussion, the functional notation of (3.19) isn't very illuminating in showing how production technology affects the supply schedule, because technology is embodied in how the relationship between the variables Q_V^s, P_V, w, and r are "tied together" by the function $g(P_V, w, r)$.

Illustration by use of a specific functional form can help make this more concrete. Suppose, for example, that the supply schedule is approximated by the linear function:

$$Q_V^s = b_0 + b_1 P_V - b_2 w - b_3 r,\ b_i \geq 0,\ i = 1, 2, 3. \tag{3.20}$$

where the b_i's are parameters. A technological advance that lets firms produce more wine for the same amount of labor and land - a shift in the production function - would be represented by a change in the b_i's. One possibility would have b_0, for example, increase. Graphically, this would be depicted as an outward parallel shift of the supply curve.

A specific functional form may also help one better grasp the expression of a supply curve in inverse form. By solving for P_V on the left-hand-side of (3.20), we get

$$P_V = \frac{-b_0}{b_1} + \frac{1}{b_1} Q_V^s + \frac{b_2}{b_1} w + \frac{b_3}{b_1} r. \tag{3.21}$$

This inverse form in general functional notation would be written as

$$P_V = g^{-1}(\overset{+}{Q_V^s}, \overset{+}{w}, \overset{+}{r}). \tag{3.22}$$

Remember, expression of either a demand or supply functional relationship in inverse form conveys no new information. It simply arranges variables in a manner more consistent with the graphical traditions of economics.

This completes our discussion of the sub-models of demand and supply. Note how each sub-model answers the question: What happens to the value of the endogenous variables in the demand sub-model, quantity demanded per unit of time, quantity supplied per unit of time in the supply sub-model - when the values of the exogenous variables change? We now combine these sub-models into an equilibrium model of demand and supply. While in each of the above sub-models the price of wine was taken as exogenous by the decision-makers, in the equilibrium model this variable will be taken as endogenous.

1.2.2 Equilibrium

The concept of "equilibrium" is that we have found a "rest point" where, in the absence of any changes in either the values of exogenous variables or in the specification of the logical interactions among these variables, the values of the endogenous variables are unchanging. For our partial-equilibrium model, this "rest point" occurs when the quantity demanded equals the quantity supplied. In symbolic notation, this *equilibrium condition* is:

$$Q_V^d = Q_V^s \tag{3.23}$$

With the addition of this equilibrium condition, the specification of our model is complete. When we have a simple listing of all the equational statements of a model, where each equation represents either a behavioral relationship, an identity, definition, or technical relationship, or an equilibrium condition, then we say the model is written in *structural form.* The structural form of our simple demand-supply model is simply the collection of the behavioral equations representing the demand function and the supply function, and the equilibrium condition that quantity demanded equals quantity supplied. For emphasis, we collect the three equations of our model here:

$$Q_V^d = f(\overset{-}{P_V}, \overset{-}{P_B}, \overset{+}{P_F}, \overset{+}{Y}) \quad (3.2)$$

$$Q_V^s = g(\overset{+}{P_V}, \overset{-}{w}, \overset{-}{r}) \quad (3.19)$$

$$Q_V^d = Q_V^s \quad (3.23)$$

Each equation spells out an economic assumption of the model, telling us something about postulated behavior of individuals-((3.2) and (3.19))- or markets-(3.23). That is, the structural model lays bare the economic framework of the model.

Our goal, though, is to use the model to answer the canonical questions we posed earlier: what is the relationship between the values of the exogenous variables and the values of the endogenous variables? To this end, we now develop various ways that we can answer this question via a solution of the model.

1.2.3 Solving the model without graphs

A strategy or "protocol" often used is the substitution of behavioral relationships into the equilibrium condition. Substitution of equation (3.1), the demand schedule, and equation (3.18), the supply schedule, into equilibrium condition (3.23) tells us implicitly what price is necessary to equate quantity demanded to quantity supplied:

$$f(P_V; P_B, P_F, Y) = g(P_V; w, r) \quad (3.24)$$

We separate the arguments in each function by a semi-colon to denote that the variables to the right of the semi-colon in each function are exogenous. This emphasizes that (3.24) is a single equation in which the only endogenous variable is P_V.

We say that this substitution tells us *implicitly* the equilibrium price that clears the market because (3.24) is an example of an implicit function. Up until now, we have encountered mostly functions of the form $y = f(x)$. In equations like this, the variable y simply appears by itself on the left hand side of the equation, and expressions which only involve the variable x appear on the right hand side. That is, there is no intermingling of y and x on either side of the equality sign, and y appears as the variable itself and not as some function of the variable such as, for example, y^2 , or perhaps $y - 6$. We say that such an equation tells us *explicitly* what the value is of y for any permissible value of x. For example, equation 3.14.ii , replicated here, expresses P_V *explicitly* as a function of Q_V^d:

$$P_V = \frac{1}{4Q_V^d}.$$

An implicit function, on the other hand, may have variables intertwined and showing up on either side of the equality signs. They are equations like

$$\begin{aligned} x + y &= 1; \\ y^2 &= x; \\ x^2 + xy + y^2 &= 3 \end{aligned}$$

As another example, we could multiply both sides of equation 3.14.ii by Q_V^d so as to turn it into an implicit function:

$$P_V Q_V^d = \frac{1}{4}$$

Of course, this also means that we can turn this *implicit* equation into an *explicit* equation by multiplying both sides by $\frac{1}{Q_V^d}$. For many implicit functions, this transformation from implicit to explicit function can be done. Of most interest to us is that the functions we will encounter in most economic problems can in principle be transformed in this way.

For the problem at hand, the implicit equation (3.24) can be transformed into an explicit equation with P_V on the left-hand-side of the equality sign and all other variables on the right. When we do this transformation, we often describe this by saying that the equilibrium price of wine can be "solved out" of equation (3.24) and expressed as a function of just the exogenous variables:

$$\widehat{P}_V = F(P_B, P_F, w, r). \tag{3.25}$$

This expression for $\widehat{P}_V$ can be substituted into either the demand or the supply function to yield an equation that expresses the equilibrium quantity of wine produced or consumed per unit of time as a function of just the exogenous variables. For example, if we substituted (3.25) into the supply function (3.18) in place of P_V , we would have

$$\widehat{Q_V^s} = g(F(P_B, P_F, w, r), w, r)$$

which we could write as:

$$\widehat{Q_V^s} = G(P_B, P_F, w, r). \tag{3.26}$$

Finally, because of the equilibrium condition,

$$\widehat{Q_V^s} = \widehat{Q_V^d} \equiv \widehat{Q}_V = G(P_B, P_F, w, r). \tag{3.27}$$

where we use the "defined as" binary relation symbol to emphasize that the equilibrium quantity per unit of time is not a different value for the demand

and supply functions. Note that we denote equilibrium values of P_V and Q_V by putting a "hat" over the symbol for the variable. This helps keep clear that when we write P_V and Q_V in the demand and supply functions (without a "hat"), they are not variables with uniquely determined values. The equilibrium values of these variables, denoted by the "hat" over them, though, are the particular values that simultaneously satisfy all the equations of the model.

"Solving out" may not yet be clear to someone without some mathematical sophistication. Again, use of specific functional forms helps make this clearer. For purposes of exposition, consider the examples of linear functional forms for demand and supply. Substitution of these equations into (3.24) and use of the tools of ordinary high school algebra yield the explicit function for the equilibrium price:

$$\widehat{P}_V = \frac{1}{a_1 + b_1}\{a_0 - b_0 - a_2 P_B + a_3 P_F + a_4 Y + b_2 w + b_3 r\}. \tag{3.28}$$

Now the equilibrium quantity bought and sold can be solved explicitly by substituting (3.28) into either the demand or supply function, and applying again the tools of ordinary high school algebra:

$$\widehat{Q}_V = \frac{1}{a_1 + b_1}\{a_0 b_1 + a_1 b_0 - a_2 b_1 P_B + a_3 b_1 P_F + a_4 b_1 Y - a_1 b_2 w - a_1 b_3 r\} \tag{3.29}$$

where again the "hat" identifies an equilibrium value. Equations (3.28) and (3.29) are ideally suited to answering the canonical question that we ask of models: what are the values of the endogenous variables for given values of the exogenous variables. If we are given values for the parameters - the a_i's and b_i's - and are given values for each of the exogenous variables - P_B, P_F, Y, w, r - then we can immediately compute the equilibrium values of P_V and Q_V.

Even with knowledge of a specific functional form such as above, we may not know the magnitudes of the various parameters. That is, all we might know about the a_i's and b_i's is that they are non-negative numbers. Clearly, we cannot ascertain exact numerical values of the In this case the form of the canonical question that we would ask would be: In what directions-larger or smaller- do the values of the endogenous variables move when there is a qualitative change,i.e., an increase of decrease of unspecified magnitude, in the value of an exogenous variable. Because all the parameters were specified as non-negative numbers, we can immediately "sign" these changes by inspection. For example, if P_B were to increase, ceterus paribus, then the value of $\widehat{P}_V$ would decrease because the coefficient on P_B in equation (3.27), namely $\frac{-a_2}{a_1+b_1}$, is negative. The value of $\widehat{Q}_V$ would also decrease because the coefficient on P_B in equation (3.28), namely $\frac{-a_1 b_1}{a_1+b_1}$, is negative.

Can we answer these questions even if we don't have specific functional forms for our demand and supply functions? For this model, we can. What will allow us to do this is the *behavioral* assumptions about which exogenous variables affect demand and which ones affect supply. We emphasize behavioral in the preceding sentence to make clear these are assumptions that are manifested in structural equations of the model and not in equations such as (3.25) and (3.26) or their specific functional form counterparts (3.28) and (3.29). Equation (3.25) and (3.26) or any of their counterparts that arise from use of a specific functional form of specific example are known as *reduced form* equations. Before seeing how we can answer the canonical questions, let us digress briefly to spell out the relationship between structural and reduced form.

To move from the structural form to the reduced form, we repeatedly substituted structural equational statements about one or another endogenous variables for that endogenous variable in another structural equation. In the above model, we substituted the right-hand side of the structural equation that had Q_V^d on the left-hand-side (equation 3.2) and we substituted the right-hand side of the structural equation that had Q_V^s on the left-hand side (equation 3.19), into the structural equation that imposed the equilibrium condition (equation 3.23). We then substituted the implicit function for P_V that arose out of this first step into the left-hand side of the structural equation for supply. These mathematical manipulations transformed the structural model into the reduced from.

For any model, when this rearrangement is completed, there is one reduced-form equation for each endogenous variable; such an equation has one and only one endogenous variable on the left-hand-side of the equation, with any number of exogenous variables on the right-hand-side. Remember, though, the right-hand-side contains no endogenous variables.

To get from a structural model to the reduced form, the structural model needs to have as many equations as there are endogenous variables. This equality of numbers of structural equations and endogenous variables is in fact a necessary condition for any structural model to be capable of answering the key question we ask of models: what happens to the values of endogenous variables when the value of an exogenous variable changes? This means that counting equations and endogenous variables provides a diagnostic check of whether or not a structural model is "well-specific", that is, whether or not it can answer our canonical question. As an example, in our demand-supply model, there are three structural equations, namely the demand function, the supply function, and the equilibrium condition, and three endogenous variables: the quantity demanded (Q_V^d), the quantity supplied (Q_V^s), and the price of wine (P_V). Note again that there is no logical restriction on the number of exogenous variables.

Now, let us return to how we go about answering the canonical question. The first step of the transformation of the structural model into the reduced form was an equating of demand and supply functions to each other. This was equation 3.24, reproduced here for convenience:

$$\overbrace{f(\overset{-}{P_V}; \overset{-}{P_B}, \overset{+}{P_F}, \overset{+}{Y})}^{Q_V^d} = \overbrace{g(\overset{+}{P_V}; \overset{-}{w}, \overset{-}{r})}^{Q_V^s} \tag{3.24}$$

The key feature of this equation is that the left-hand side has the single endogenous variable P_V and the exogenous variables P_B, P_F, and Y, while the right-hand side has the same single endogenous variable P_V but the *different* exogenous variables w and r. This reflects our behavioral assumptions about what exogenous variables affect the quantity demanded and quantity supplied, respectively.

Carry out the thought experiment of contemplating a ceterus paribus increase in the value of one of the exogenous variables, say Y, from an initial position of equilibrium. What happens to P_V? There are only three possibilities: It must either stay the same, decrease, or increase. We will trace out the implications of each of these possibilities and find that only one is possible. Because these three possibilities exhaust all cases, the one that is possible is also the correct answer.

Suppose P_V remained the same. This would mean the right-hand side of (3.24), that is, Q_V^s, remains the same number. This of course requires that the left-hand side-Q_V^d- remains that same number. But if Y increases, ceterus paribus, and P_V remains unchanged, then Q_V^d must have gone up. But if Q_V^d increased in value, it cannot still be equal to the unchanged value of Q_V^s. We have a contradiction. This eliminates the possibility that P_V remains unchanged in the face of an increase in Y.

Now suppose P_V decreased. If this happens, the right hand side of (3.24) requires that Q_V^s also decreases. (This is just the behavioral assumption that, ceterus paribus, quantity supplied increases with an increase in its own price.) If Q_V^s decreases, then Q_V^d must also decrease. But *behaviorally*, a decrease in P_V, ceterus paribus, increases the value of the variable Q_V^d. By assumption, the only other variable to change value is Y, and behaviorally an increase in Y also increases the value of the variable Q_V^d. Hence, if both P_V and Y increase, ceterus paribus, then Q_V^s must have increased. But this means that it cannot be equal to Q_V^s. We have a contradiction. This eliminates the second possibility that an increase in Y could lead to a decrease in P_V.

Hence the only possibility left is that P_V increased, so this must be the correct answer. (Remember, one of the three possibilities had to occur.) A diagnostic check on our logic can be done by seeing what happened to Q_V^s from such a rise in price. Our behavioral assumption tells us that Q_V^s must have risen in value. This means that Q_V^d must also have increased in value. This is consistent with an increase in both P_V and Y.

Note that if we had simply started our analysis with the possibility that there was an increase in P_V , all we would have proved was that this possibility could not be ruled out. But because we ruled out all other possibilities, this last possibility must in fact be what had to have happened.

Knowing that P_V must have increased in response to a ceterus paribus increase in Y, we can now figure out what is the response of Q_V. Because the exogenous variables in the supply function,i.e., the right-hand side of (3.24), are, by assumption, unchanged, the behavioral assumption that Q_V^s increases if P_V increases insures that the equilibrium value of Q_V (which must equal both the quantity demanded, the value of the left-hand side of (3.24), and the quantity supplied, the value of the right-hand side) has increased.

To sum up, we have been able to infer the answer to the "directional"form of our canonical questions about the effects of an increase in Y on the values of the endogenous variables Q_V and P_V even without knowledge of specific functional forms. For many people, though, a graphical approach to answering these questions is easier to understand.

Graphical solution Graphically, market equilibrium is depicted by superimposing on one graph both the demand and the supply curve. Remember, a demand function is a set of points that constitutes a downward-sloping line-not necessarily straight- in the $P_V - Q_V^d$ plane, and a supply function is a set of points that constitutes an upward-sloping line-again not necessarily straight-in the $P_V - Q_V^s$ plane. In terms of inverse demand and supply functions, these curves have the same slopes but are in the $Q_V - P_V$ plane. That is, the vertical axis measures price and the horizontal axis measures quantity. The intersection of these curves depicts the equilibrium price and the equilibrium quantities bought and sold per unit of time.. Notice that the equilibrium is an ordered pair, represented by a point in the plane. This is depicted in Figure 3.13 as the point $(1, 1)$. The labels on the horizontal axis are both Q_V^d and Q_V^s -one for the demand function and one for the supply function. The point on the horizontal axis where $Q_V^d = Q_V^s$ is that value of Q where the quantity demanded equals the quantity supplied. Associated with this equilibrium value of Q_V is the equilibrium value of P_V.

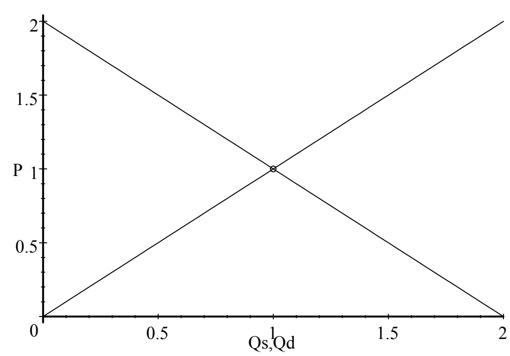

Figure 3.13: Equilibrium

Although the graph uses linear functions to depict this equilibrium, all the information we need to know about the demand and supply curves in order to depict an equilibrium is that they intersect at some point in the positive quadrant of the $Q_V - P_V$ plane. We don't need to know a specific functional form or specific parameter values. Furthermore, without knowing a specific functional form or specific parameter values, we can further characterize the relationship between exogenous and endogenous variables based solely on our qualitative behavioral knowledge captured be the assumed algebraic signs over the arguments in the behavioral equations that constitute the structural model. This is done by applying the "curve-shifting" techniques introduced in our analysis of demand and supply.

The key feature of our model that lets us easily use the diagrammatic analysis to characterize the effects of changes in values of exogenous variables on the values of endogenous variables is the lack of overlap of exogenous variables in the demand and supply behavioral relationships. Because of this, any of our ceterus paribus thought experiments in which we assume a change in the value of just a single exogenous variable results in a shift of only one curve: either the demand (or inverse demand) curve or the supply (or inverse supply) curve. Hence, a change in, say, income, shifts 'out" or "up" the inverse demand curve (or, equivalently, shifts us to a member of the family of inverse demand curves that is farther "out" or "up" from the original member). The inverse supply curve, though, remains at its initial placement in the $Q_V - P_V$ plane. Consequently, the new equilibrium point is a new pair (Q_V, P_V) that we can think of as representing a movement along the inverse supply curve resulting from a shift in the inverse demand curve. Figure 3.14 depicts such a shift of the inverse demand curve and the associated movement of the equilibrium pair along the inverse supply curve. In the figure, the new inverse demand curve is depicted as a red line. Clearly, the new equilibrium pair has a higher value for both Q_V and P_V.

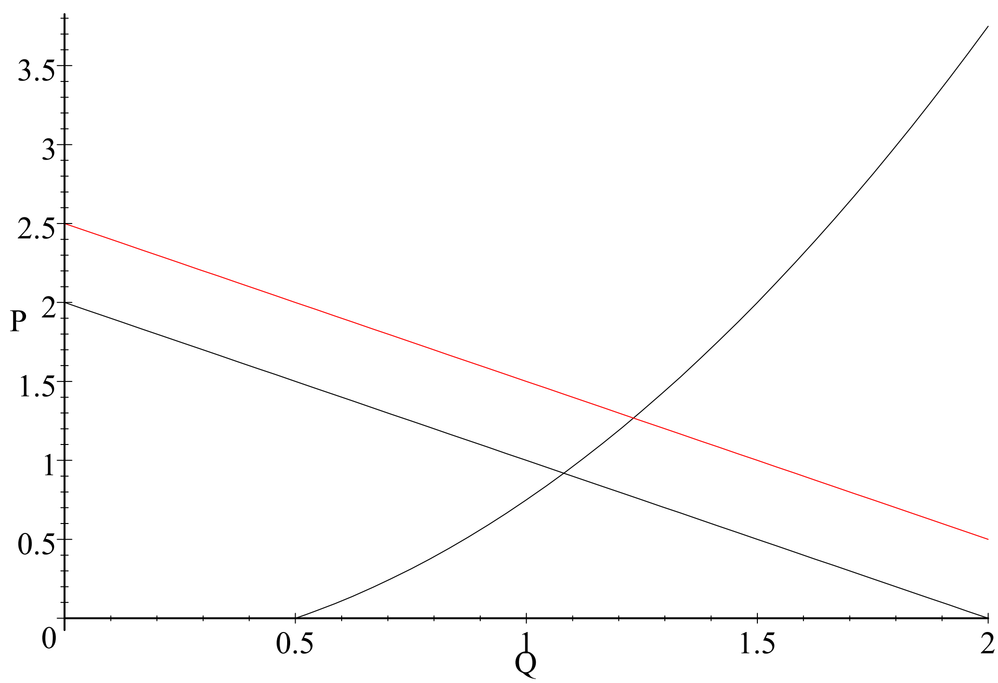

Figure 3.14: Equilibrium after an increase in income

Other ceterus paribus thought experiments follow the same logic. For example, if w were to increase, this would shift up the inverse supply curve while leaving the inverse demand curve unchanged. The new equilibrium could be described as a new pair (Q_V, P_V) that represents a movement along the inverse demand curve resulting from a shift in the supply curve. This new pair would have a higher value of P_V and a lower value of Q_V.

The value of having specified graphical relationships in which only one curve shifts in response to a change in an exogenous variable is that once the direction of the shift is known, the change in the equilibrium values of the endogenous variables is obvious. For our simple partial-equilibrium model of demand and supply, the specification of our behavioral relationships led us naturally to such a diagram. In more complicated models, we frequently manipulate the structural model solely to be able to depict the model in terms of this type of graph. Such manipulations do not reflect some deep economic insight, but are rather a strategy used to display relationships in this useful way.

Some Further Questions We have now provided a careful statement of the partial-equilibrium demand-supply model familiar from an economic-principles course. While providing a review of this basic model, the primary purpose has been to illustrate the generic structure of economic models and to familiarize the reader with the terminology of these models. Two interrelated questions remain. First, how do we evaluate this model? Is it a ”good” model, and what would we mean by that? Second, are there questions not adequately addressed by this model that suggest the need for a more refined model?

Evaluation of Models What makes a ”good” model? One answer to this question is given by how close a model hews to the five ”epistemic virtues” that were spelled out in Chapter Two: predictive ability, internal coherence and external consistency, unification, fertility, and simplicity. A less encompassing but still useful answer is that a good model also helps us understand the problem at hand. At a very preliminary stage of investigation of a problem, a good model can be one that simply helps us organize thought about the problem. For many questions, the simple demand-supply model just presented is a ”good” model.

Unanswered Questions For some of the most interesting questions of international trade, though, this model is inadequate. For one thing, what is needed to grapple with the idea of what economists mean by ”gains from trade” is a model that looks behind the demand and supply curves and provides a fuller description of what is meant by tastes and technology, and how these notions constrain behavior. For another, much of the subject matter of international economics requires an understanding of the simultaneous interactions between the parts of all of the markets of an economy. Such a general equilibrium model has two important advantages over partial-equilibrium models for the purposes of understanding international trade issues. First, some of the most prominent arguments about the effects of trade policy concerns questions about whether jobs are created or destroyed. Remember, for example, the discussions in the introductory chapter about the loss of jobs in the textile industry arising from rising foreign imports, or the expected loss of jobs in "big steel" unless tariffs were imposed. General equilibrium models emphasize and make clear that on net, jobs are not "lost" in response to changing economic conditions, but are redistributed from one sector to another. Once one develops the habit of thinking in terms of general equilibrium, one quickly recognizes the less-emphasized related effects of job creation in other sectors whenever the "jobs" discussions about trade policy is broached. This is not to say that there are not real problems and hardships associated with the economic dislocations brought about by participation in the international economy, but rather that there are also opportunities not as readily apparent to someone not trained in general-equilibrium thinking.

Such general-equilibrium habits of thought also help one bring to mind more quickly analogous situations that help one think about trade policy effects. For example, other kinds of economic changes such as changes in technology have the same types of dislocation effects as does international trade. The development of word processors caused a loss of jobs in the typewriter manufacturing industry, but led to an expansion of a variety of other industries. This situation is analogous to the "textile" question in that both of these situations had the same generic effects on employment: jobs in some industries were lost, and jobs in other industries were gained.

The jobs issue is just one example of a general phenomenon highlighted by general-equilibrium models: the substitutability of resources within an economy. Recall from the first chapter the discussion about the popular view of "price gouging" by hoteliers in the research triangle area of North Carolina in anticipation of soon-to-be-held special Olympics. This view, we argued, was flawed because it failed to understand the ability of producers in an economy to substitute resources in the production of goods and services, and the ability of consumers in an economy to substitute one good for another in the face of changing incentives. This feature of *substitutability*, as one of the few key elements of the way economist's think about the world, is really only fully understood and appreciated within the context of a general equilibrium model.

We now move on to develop three key models for an understanding of international economics. First, we develop the simplest possible general equilibrium model we can envisage, one in which there are no production decisions. Such an economy, populated only by consumers but not producers, allows us to lay bare the basic novel features of general equilibrium models and some key features of international trade models. It also allows us to introduce the concept of "gains from trade" in its most pure and uncomplicated form.

Second, we introduce production into the general equilibrium framework. Because so much of international economics concerns the effects of changes in the mix of products produced by the factors of production available to the economy as a whole, knowledge of this model is essential to an understanding of both international trade and open-economy macroeconomics.

Finally, we analyze the demand and supply of money. Knowledge of how economists model this feature of the modern economy permits us to understand the key distinction between real and nominal variable.