MATHEMATICAL APPENDIX

In this Appendix we will provide a brief review of some of the mathematical concepts that are used in the text. This material is meant to serve as a reminder of the definitions of various terms used in the text. It is emphatically not a tutorial in mathematics. The definitions given will generally be the simplest, not the most rigorous.

A.1 Functions

A **function** is a rule that describes a relationship between numbers. For each number x, a function assigns a *unique* number y according to some rule. Thus a function can be indicated by describing the rule, as "take a number and square it," or "take a number and multiply it by 2," and so on. We write these particular functions as $y = x^2$, $y = 2x$. Functions are sometimes referred to as **transformations**.

Often we want to indicate that some variable y depends on some other variable x, but we don't know the specific algebraic relationship between the two variables. In this case we write $y = f(x)$, which should be interpreted as saying that the variable y depends on x according to the rule f.

Given a function $y = f(x)$, the number x is often called the **independent variable**, and the number y is often called the **dependent variable**.

The idea is that x varies independently, but the value of y *depends* on the value of x.

Often some variable y depends on several other variables x_1, x_2, and so on, so we write $y = f(x_1, x_2)$ to indicate that both variables together determine the value of y.

A.2 Graphs

A **graph** of a function depicts the behavior of a function pictorially. Figure A.1 shows two graphs of functions. In mathematics the independent variable is usually depicted on the horizontal axis, and the dependent variable is depicted on the vertical axis. The graph then indicates the relationship between the independent and the dependent variables.

However, in economics it is common to graph functions with the independent variable on the vertical axis and the dependent variable on the horizontal axis. Demand functions, for example, are usually depicted with the price on the vertical axis and the amount demanded on the horizontal axis.

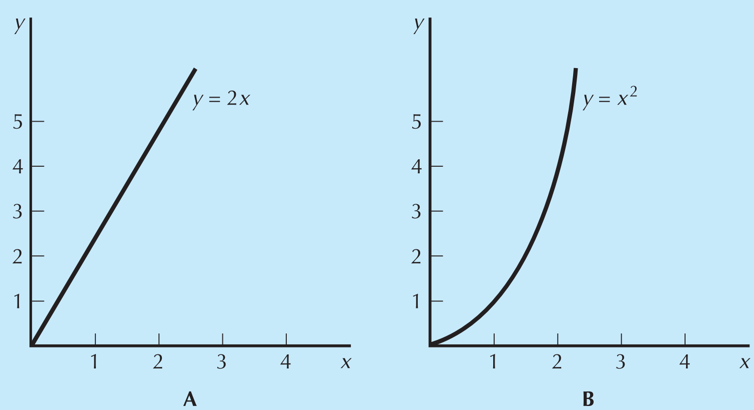

Figure A.1 **Graphs of functions**. Panel A denotes the graph of $y = 2x$, and panel B denotes the graph of $y = x^2$.

A.3 Properties of Functions

A **continuous function** is one that can be drawn without lifting a pencil from the paper: there are no jumps in a continuous function. A **smooth**

function is one that has no "kinks" or corners. A **monotonic** function is one that always increases or always decreases; a **positive monotonic function** always increases as x increases, while a **negative monotonic function** always decreases as x increases.

A.4 Inverse Functions

Recall that a function has the property that for each value of x there is a unique value of y associated with it and that a monotonic function is one that is always increasing or always decreasing. This implies that for a monotonic function there will be a unique value of x associated with each value of y.

We call the function that relates x to y in this way an **inverse function**. If you are given y as a function of x, you can calculate the inverse function just by solving for x as a function of y. If $y = 2x$, then the inverse function is $x = y/2$. If $y = x^2$, then there is no inverse function; given any y, both $x = +\sqrt{y}$ and $x = -\sqrt{y}$ have the property that their square is equal to y. Thus there is not a *unique* value of x associated with each value of y, as is required by the definition of a function.

A.5 Equations and Identities

An **equation** asks when a function is equal to some particular number. Examples of equations are

$$\begin{aligned} 2x &= 8 \\ x^2 &= 9 \\ f(x) &= 0. \end{aligned}$$

The **solution** to an equation is a value of x that satisfies the equation. The first equation has a solution of $x = 4$. The second equation has two solutions, $x = 3$ and $x = -3$. The third equation is just a general equation. We don't know its solution until we know the actual rule that f stands for, but we can denote its solution by x^*. This simply means that x^* is a number such that $f(x^*) = 0$. We say that x^* **satisfies** the equation $f(x) = 0$.

An **identity** is a relationship between variables that holds for *all* values of the variables. Here are some examples of identities:

$$\begin{aligned} (x+y)^2 &\equiv x^2 + 2xy + y^2 \\ 2(x+1) &\equiv 2x + 2. \end{aligned}$$

The special symbol $\equiv$ means that the left-hand side and the right-hand side are equal for *all* values of the variables. An equation only holds for some values of the variables, whereas an identity is true for all values of the variables. Often an identity is true by the definition of the terms involved.

A.6 Linear Functions

A **linear function** is a function of the form

$$y = ax + b,$$

where a and b are constants. Examples of linear functions are

$$y = 2x + 3$$
$$y = x - 99.$$

Strictly speaking, a function of the form $y = ax + b$ should be called an **affine function**, and only functions of the form $y = ax$ should be called linear functions. However, we will not insist on this distinction.

Linear functions can also be expressed implicitly in forms like $ax+by = c$. In such a case, we often like to solve for y as a function of x to convert this to the "standard" form:

$$y = \frac{c}{b} - \frac{a}{b}x.$$

A.7 Changes and Rates of Change

The notation Δx is read as "the change in x." It does *not* mean Δ times x. If x changes from x^* to x^{**}, then the change in x is just

$$\Delta x = x^{**} - x^*.$$

We can also write

$$x^{**} = x^* + \Delta x$$

to indicate that x^{**} is x^* plus a change in x.

Typically Δx will refer to a *small* change in x. We sometimes express this by saying that Δx represents a **marginal change**.

A **rate of change** is the ratio of two changes. If y is a function of x given by $y = f(x)$, then the rate of change of y with respect to x is denoted by

$$\frac{\Delta y}{\Delta x} = \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

The rate of change measures how y changes as x changes.

A linear function has the property that the rate of change of y with respect to x is constant. To prove this, note that if $y = a + bx$, then

$$\frac{\Delta y}{\Delta x} = \frac{a + b(x + \Delta x) - a - bx}{\Delta x} = \frac{b\Delta x}{\Delta x} = b.$$

For nonlinear functions, the rate of change of the function will depend on the value of x. Consider, for example, the function $y = x^2$. For this function

$$\frac{\Delta y}{\Delta x} = \frac{(x + \Delta x)^2 - x^2}{\Delta x} = \frac{x^2 + 2x\Delta x + (\Delta x)^2 - x^2}{\Delta x} = 2x + \Delta x.$$

Here the rate of change from x to $x + \Delta x$ depends on the value of x and on the size of the change, Δx. But if we consider very small changes in x, Δx will be nearly zero, so the rate of change of y with respect to x will be approximately $2x$.

A.8 Slopes and Intercepts

The rate of change of a function can be interpreted graphically as the **slope** of the function. In Figure A.2A we have depicted a linear function $y = -2x + 4$. The **vertical intercept** of this function is the value of y when $x = 0$, which is $y = 4$. The **horizontal intercept** is the value of x when $y = 0$, which is $x = 2$. The slope of the function is the rate of change of y as x changes. In this case, the slope of the function is -2.

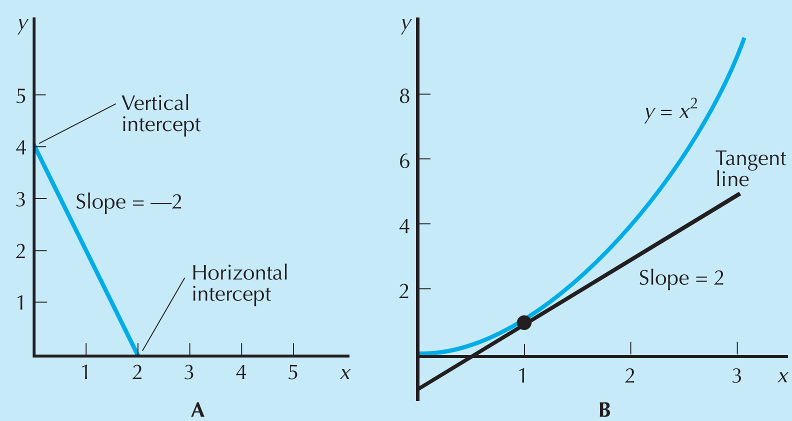

Figure A.2 **Slopes and intercepts.** Panel A depicts the function $y = -2x + 4$, and panel B depicts the function $y = x^2$.

In general, if a linear function has the form $y = ax + b$, the vertical intercept will be $y^* = b$ and the horizontal intercept will be $x^* = -b/a$. If a linear function is expressed in the form

$$a_1x_1 + a_2x_2 = c,$$

then the horizontal intercept will be the value of x_1 when $x_2 = 0$, which is $x_1^* = c/a_1$, and the vertical intercept will occur when $x_1 = 0$, which means $x_2^* = c/a_2$. The slope of this function is $-a_1/a_2$.

A nonlinear function has the property that its slope changes as x changes. A **tangent** to a function at some point x is a linear function that has the same slope. In Figure A.2B we have depicted the function x^2 and the tangent line at $x = 1$.

If y increases whenever x increases, then Δy will always have the same sign as Δx, so that the slope of the function will be positive. If on the other hand y decreases when x increases, or y increases when x decreases, Δy and Δx will have opposite signs, so that the slope of the function will be negative.

A.9 Absolute Values and Logarithms

The **absolute value** of a number is a function $f(x)$ defined by the following rule:

$$f(x) = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0. \end{cases}$$

Thus the absolute value of a number can be found by dropping the sign of the number. The absolute value function is usually written as $|x|$.

The (natural) **logarithm** or **log** of x describes a particular function of x, which we write as $y = \ln x$ or $y = \ln(x)$. The logarithm function is the unique function that has the properties

$$\ln(xy) = \ln(x) + \ln(y)$$

for all positive numbers x and y and

$$\ln(e) = 1.$$

(In this last equation, e is the base of natural logarithms which is equal to $2.7183\ldots$) In words, the log of the product of two numbers is the sum of the individual logs. This property implies another important property of logarithms:

$$\ln(x^y) = y\ln(x),$$

which says that the log of x raised to the power y is equal to y times the log of x.

A.10 Derivatives

The **derivative** of a function $y = f(x)$ is defined to be

$$\frac{df(x)}{dx} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

In words, the derivative is the limit of the rate of change of y with respect to x as the change in x goes to zero. The derivative gives precise meaning to the phrase "the rate of change of y with respect to x for small changes in x." The derivative of $f(x)$ with respect to x is also denoted by $f'(x)$.

We have already seen that the rate of change of a linear function $y = ax + b$ is constant. Thus for this linear function

$$\frac{df(x)}{dx} = a.$$

For a nonlinear function the rate of change of y with respect to x will usually depend on x. We saw that in the case of $f(x) = x^2$, we had $\Delta y/\Delta x = 2x + \Delta x$. Applying the definition of the derivative

$$\frac{df(x)}{dx} = \lim_{\Delta x \to 0} 2x + \Delta x = 2x.$$

Thus the derivative of x^2 with respect to x is $2x$.

It can be shown by more advanced methods that if $y = \ln x$, then

$$\frac{df(x)}{dx} = \frac{1}{x}.$$

A.11 Second Derivatives

The **second derivative** of a function is the derivative of the derivative of that function. If $y = f(x)$, the second derivative of $f(x)$ with respect to x is written as $d^2 f(x)/dx^2$ or $f''(x)$. We know that

$$\frac{d(2x)}{dx} = 2$$
$$\frac{d(x^2)}{dx} = 2x.$$

Thus

$$\frac{d^2(2x)}{dx^2} = \frac{d(2)}{dx} = 0$$
$$\frac{d^2(x^2)}{dx^2} = \frac{d(2x)}{dx} = 2.$$

The second derivative measures the curvature of a function. A function with a negative second derivative at some point is concave near that point; its slope is decreasing. A function with a positive second derivative at a point is convex near that point; its slope is increasing. A function with a zero second derivative at a point is flat near that point.

A.12 The Product Rule and the Chain Rule

Suppose that $g(x)$ and $h(x)$ are both functions of x. We can define the function $f(x)$ that represents their product by $f(x) = g(x)h(x)$. Then the derivative of $f(x)$ is given by

$$\frac{df(x)}{dx} = g(x)\frac{dh(x)}{dx} + h(x)\frac{dg(x)}{dx}.$$

Given two functions $y = g(x)$ and $z = h(y)$, the **composite function** is

$$f(x) = h(g(x)).$$

For example, if $g(x) = x^2$ and $h(y) = 2y + 3$, then the composite function is

$$f(x) = 2x^2 + 3.$$

The **chain rule** says that the derivative of a composite function, $f(x)$, with respect to x is given by

$$\frac{df(x)}{dx} = \frac{dh(y)}{dy}\frac{dg(x)}{dx}.$$

In our example, $dh(y)/dy = 2$, and $dg(x)/dx = 2x$, so the chain rule says that $df(x)/dx = 2 \times 2x = 4x$. Direct calculation verifies that this is the derivative of the function $f(x) = 2x^2 + 3$.

A.13 Partial Derivatives

Suppose that y depends on both x_1 and x_2, so that $y = f(x_1, x_2)$. Then the **partial derivative** of $f(x_1, x_2)$ with respect to x_1 is defined by

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = \lim_{\Delta x_1 \to 0} \frac{f(x_1 + \Delta x_1, x_2) - f(x_1, x_2)}{\Delta x_1}.$$

The partial derivative of $f(x_1, x_2)$ with respect to x_1 is just the derivative of the function with respect to x_1, *holding x_2 fixed*. Similarly, the partial derivative with respect to x_2 is

$$\frac{\partial f(x_1, x_2)}{\partial x_2} = \lim_{\Delta x_2 \to 0} \frac{f(x_1, x_2 + \Delta x_2) - f(x_1, x_2)}{\Delta x_2}.$$

Partial derivatives have exactly the same properties as ordinary derivatives; only the name has been changed to protect the innocent (that is, people who haven't seen the ∂ symbol).

In particular, partial derivatives obey the chain rule, but with an extra twist. Suppose that x_1 and x_2 both depend on some variable t and that we define the function $g(t)$ by

$$g(t) = f(x_1(t), x_2(t)).$$

Then the derivative of $g(t)$ with respect to t is given by

$$\frac{dg(t)}{dt} = \frac{\partial f(x_1, x_2)}{\partial x_1} \frac{dx_1(t)}{dt} + \frac{\partial f(x_1, x_2)}{\partial x_2} \frac{dx_2(t)}{dt}.$$

When t changes, it affects both $x_1(t)$ and $x_2(t)$. Therefore, we need to calculate the derivative of $f(x_1, x_2)$ with respect to each of those changes.

A.14 Optimization

If $y = f(x)$, then $f(x)$ achieves a **maximum** at x^* if $f(x^*) \geq f(x)$ for all x. It can be shown that if $f(x)$ is a smooth function that achieves its maximum value at x^*, then

$$\frac{df(x^*)}{dx} = 0$$
$$\frac{d^2 f(x^*)}{dx^2} \leq 0.$$

These expressions are referred to as the **first-order condition** and the **second-order condition** for a maximum. The first-order condition says that the function is flat at x^*, while the second-order condition says that the function is concave near x^*. Clearly both of these properties have to hold if x^* is indeed a maximum.

We say that $f(x)$ achieves its **minimum** value at x^* if $f(x^*) \leq f(x)$ for all x. If $f(x)$ is a smooth function that achieves its minimum at x^*, then

$$\frac{df(x^*)}{dx} = 0$$
$$\frac{d^2 f(x^*)}{dx^2} \geq 0.$$

The first-order condition again says that the function is flat at x^*, while the second-order condition now says that the function is convex near x^*.

If $y = f(x_1, x_2)$ is a smooth function that achieves its maximum or minimum at some point (x_1^*, x_2^*), then we must satisfy

$$\frac{\partial f(x_1^*, x_2^*)}{\partial x_1} = 0$$
$$\frac{\partial f(x_1^*, x_2^*)}{\partial x_2} = 0.$$

These are referred to as the **first-order conditions**. There are also second-order conditions for this problem, but they are more difficult to describe.