

4. DERIVATIVES AND DIFFERENTIATION RULES

J.P. Bastos*

EPH 1100: Analytical Tools

Reading: Hoy et al., §5.2–5.3, §5.4

- **§5.2:** Definition of the derivative. Read alongside these notes.
- **§5.3:** Conditions for differentiability. Skim; the key takeaway is Theorem 5.1.
- **§5.4:** Rules of Differentiation. Work through the examples.

1. From Geometry to Algebra

In Class 3, we defined the slope of a curve at a point as the limit of a sequence of secant slopes:

$$m^* = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

That limit, when it exists, is the **derivative** of f at x . We gave it a geometric interpretation (the slope of the tangent line) and connected it to the idea of an instantaneous rate of change.

This session makes the concept operational. Instead of computing limits from scratch every time, we derive a set of differentiation rules that let us find derivatives quickly for the functions that arise in economics. The payoff is **marginal analysis**: the ability to ask, precisely, how much an output changes for a small change in an input.

2. The Derivative

The derivative of a function at a point captures how fast the function is changing there (the instantaneous rate of change, or the slope of the curve at that exact point).

Definition. The **derivative** of $y = f(x)$ at a point x is

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

provided this limit exists. Alternative notations: $f'(x)$, $\frac{dy}{dx}$, $\frac{df}{dx}$.

* Assistant Professor of Economics, University of Austin (UTAX). jbastos@uastin.org

The derivative is itself a function: it assigns to each x the slope of f at that point. For a linear function $f(x) = mx + b$, the slope is the same everywhere ($f'(x) = m$), so the derivative is a constant. For a nonlinear function like $f(x) = x^2$, the slope varies with x and the derivative is a new function ($f'(x) = 2x$). That is, the slope of $f(x) = x^2$ changes with x : it is -2 at $x = -1$ and $+2$ at $x = 1$. The derivative $f'(x) = 2x$ encodes this.

Definition.

If $f'(x) > 0$ at some point, then the function is **increasing**;

If $f'(x) < 0$ at some point, then the function is **decreasing**.

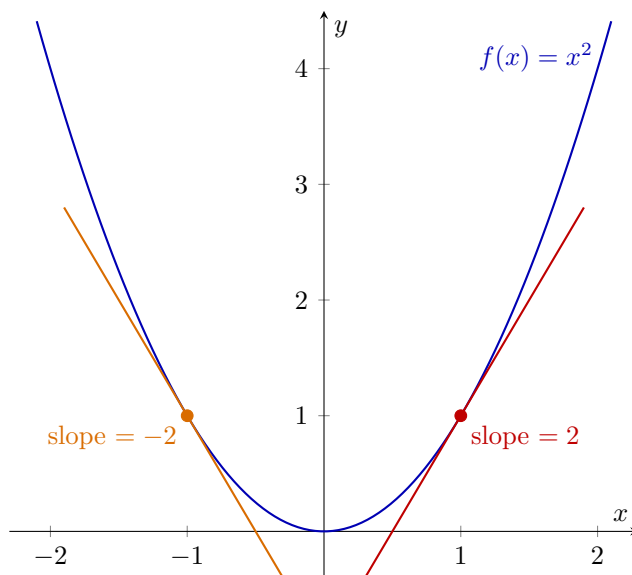


Figure 1: Tangent lines to $f(x) = x^2$.

Notes: At $x = 1$ the tangent has slope $2 = f'(1)$; at $x = -1$ the slope is $-2 = f'(-1)$. The derivative gives the exact slope at each point on the curve.

Differentiability and continuity

A function is **differentiable** at x if the derivative exists there. Not every continuous function is differentiable: a function with a “kink” (a corner, where the left and right slopes differ) is continuous but not differentiable at the kink. The converse, however, is always true: if f is differentiable at x , then f must be continuous at x . Differentiability is a stronger requirement than continuity.¹

¹Differentiability is technically a *pointwise* concept: a function is differentiable *at a point* x if the limit defining $f'(x)$ exists there, and differentiable on an interval if it is differentiable at every point in that interval. This matters because you can sometimes compute the derivative at specific points even when no single closed-form expression $f'(x)$ is valid across the whole domain. For $f(x) = |x|$: the derivative is well-defined at every $x \neq 0$ (equal to $+1$ or -1), so $f'(3) = 1$ and $f'(-2) = -1$ are perfectly meaningful. But $f'(0)$ does not exist because the left and right limits of the difference quotient disagree. The ability to write “ $f'(x) = \dots$ ” as a formula across all x requires differentiability everywhere; the ability to evaluate f' at a specific point requires only local smoothness at that point.

The canonical example is $f(x) = |x|$, shown below. The function is continuous everywhere (no jumps or gaps) but has a sharp corner at $x = 0$. Approaching from the left, the slope is -1 ; approaching from the right, it is $+1$. Since these disagree, the limit defining $f'(0)$ does not exist.

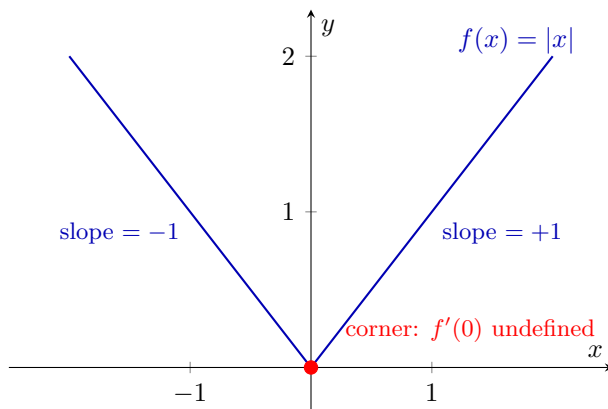


Figure 2: The absolute value function $f(x) = |x|$.

Notes: The function is continuous everywhere but has a corner at $x = 0$: the slope is -1 from the left and $+1$ from the right, so $f'(0)$ does not exist.

3. Differentiation Rules

Rather than computing every derivative from the limit definition, we use general rules derived once from that definition.

Rule 1: Constant function

$$f(x) = c \implies f'(x) = 0$$

A constant function is a horizontal line. Its slope is zero everywhere; $f(x)$ doesn't change because x doesn't change in the first place: $\Delta x = 0$. Algebraically: $f(x + \Delta x) - f(x) = f(x + 0) - f(x) = c - c = 0$, so the limit is $0/\Delta x \rightarrow 0$.

Example: A competitive firm faces a fixed market price $\bar{p}$. Its marginal revenue $MR(q) = \bar{p}$ is constant in output q , so $d(MR)/dq = 0$. The extra revenue from one more unit sold does not depend on how much the firm already sells.

Rule 2: Linear function

$$f(x) = mx + b \implies f'(x) = m$$

$f(x + \Delta x) - f(x) = m\Delta x$, so $\Delta y/\Delta x = m$ regardless of Δx . The limit is m .

Example: Linear demand: $Q^D = 100 - 2P$. Then $dQ^D/dP = -2$. Demand falls by exactly 2 units per \$1 price increase, at every price level.

Rule 3: Power function

$$f(x) = x^n \implies f'(x) = nx^{n-1} \quad (\text{for any real } n)$$

This is the most used rule in this course. It handles polynomials and a wide range of other functions (since $x^{1/2} = \sqrt{x}$, $x^{-1} = 1/x$, etc. are all power functions). Notice that Rules 2 and 3 are really the same rule. A linear function $f(x) = mx$ is a polynomial of degree 1, even if that is implicit: $f(x) = m \cdot x^1$.²

Examples:

$$\begin{aligned} f(x) = x^3 &\implies f'(x) = 3x^2 \\ f(x) = x^{1/2} &\implies f'(x) = \frac{1}{2}x^{-1/2} \\ f(x) = x^{-1} = 1/x &\implies f'(x) = -x^{-2} = -1/x^2 \end{aligned}$$

The last case explains why the slope of the rectangular hyperbola $y = 1/x$ is negative and becomes steeper (in absolute value) near the origin.

Rule 4: Constant multiple

$$g(x) = c \cdot f(x) \implies g'(x) = c \cdot f'(x)$$

Scaling the output by c scales the rate of change by c as well.

Examples:

$$\begin{aligned} g(x) = 5x^3 &\implies g'(x) = 5 \cdot 3x^2 = 15x^2 \\ g(x) = \frac{1}{2}x^4 &\implies g'(x) = \frac{1}{2} \cdot 4x^3 = 2x^3 \end{aligned}$$

Rule 5: Sum and difference

$$h(x) = f(x) \pm g(x) \implies h'(x) = f'(x) \pm g'(x)$$

Think of a firm's total cost as the sum of two independent components: labor cost $L(q)$ and material cost $M(q)$, so $TC(q) = L(q) + M(q)$. The marginal cost of producing one more unit is the extra labor plus the extra materials: the two components don't interact, so you can handle them separately and add the results. There is nothing special about this: mathematically, limits distribute over sums, $\lim(A + B) = \lim A + \lim B$, so derivatives do too.

²Applying Rule 3 to the implicit exponent 1 $f'(x) = m \cdot 1 \cdot x^{1-1} = m \cdot x^0 = m$. We state Rule 2 separately because linear functions arise constantly and the result is worth knowing directly, but it is not an independent rule.

Rule 6: Sum of n functions

Rules 4 and 5 combine immediately: for any polynomial $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, differentiate term by term:

$$f'(x) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \cdots + a_1$$

This is not really a new rule. A polynomial is a single function with $n+1$ terms, not a collection of $n+1$ separate functions that happen to be added together. You are not specifying anything beyond $f(x)$ itself. This rule is just making explicit that you can differentiate term by term in one pass.

Rule 7: Product Rule

If $h(x) = f(x) \cdot g(x)$, then:

$$h'(x) = f'(x) g(x) + f(x) g'(x)$$

Think of a rectangle with sides $f(x)$ and $g(x)$, so its area is $h(x) = f(x) \cdot g(x)$. When x increases a little, both sides grow: f expands by roughly $f'(x) \Delta x$ and g expands by roughly $g'(x) \Delta x$. The change in area comes from two new strips: one of height $g(x)$ and width $f'(x) \Delta x$ (the expansion of f , with g held fixed), and one of height $f(x)$ and width $g'(x) \Delta x$ (the expansion of g , with f held fixed). There is also a tiny corner piece at the top right, visible in Figure 3, with area $f'(x) \Delta x \cdot g'(x) \Delta x = f'(x) g'(x) (\Delta x)^2$. Why does the corner vanish but the strips don't?

First, to see this algebraically, write out all three pieces of the area change:

$$\Delta \text{Area} = \underbrace{f'(x) g(x) \Delta x}_{\text{red strip}} + \underbrace{f(x) g'(x) \Delta x}_{\text{green strip}} + \underbrace{f'(x) g'(x) (\Delta x)^2}_{\text{corner}}$$

Dividing through by Δx and taking the limit:

$$h'(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta \text{Area}}{\Delta x} = \lim_{\Delta x \rightarrow 0} \left[f'(x) g(x) + f(x) g'(x) + f'(x) g'(x) \Delta x \right]$$

The first two terms are constants with respect to the limit. The third still carries a factor of Δx : as $\Delta x \rightarrow 0$, it goes to zero regardless of how large $f'(x)$ or $g'(x)$ are. The corner contributes nothing, and $h'(x) = f'(x) g(x) + f(x) g'(x)$.

Now look at the figure and think this way: each strip has one full-size dimension and one tiny dimension: the red strip is $g(x)$ tall but only $f'(x) \Delta x$ wide; the green strip is $f(x)$ wide but only $g'(x) \Delta x$ tall. The corner is tiny in both directions. As you shrink Δx toward zero, imagine squeezing the figure: the red strip becomes a sliver, but it always stays $g(x)$ units tall; the green strip still is $f(x)$ units wide. Both have at least one dimension. The corner, however, loses width and height at the same time, so it has no dimensions. It completely vanishes.

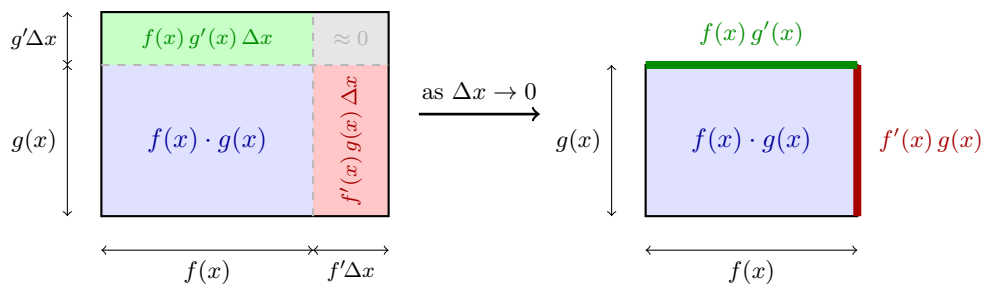


Figure 3: A visualization of the product rule

Notes: Left: for a finite $\Delta x > 0$, area growth has three pieces: two strips (red and green) and a corner (gray). Right: as $\Delta x \rightarrow 0$, all areas shrink to zero, but the strips shrink in only one dimension and collapse to one-dimensional colored border lines, while the corner shrinks in both directions and disappears entirely.

Example: Differentiate $h(x) = (3x^2 + 1)(x^3 - 2x)$.

Step 1: Define $f(x)$ and $g(x)$ and find $f'(x)$ and $g'(x)$:

$$\text{If: } f(x) = 3x^2 + 1 \quad \text{and} \quad g(x) = x^3 - 2x$$

$$\text{then: } f'(x) = 6x \quad \text{and} \quad g'(x) = 3x^2 - 2$$

Plugging into $h'(x) = f'(x)g(x) + f(x)g'(x)$:

$$h'(x) = 6x(x^3 - 2x) + (3x^2 + 1)(3x^2 - 2)$$

You could expand this, but often leaving it in this factored form is sufficient.

Application: marginal revenue

A monopolist's total revenue is $TR(q) = p(q) \cdot q$, where $p(q)$ is the inverse demand function (price expressed as a function of quantity sold). By the product rule:

$$MR(q) = TR'(q) = p'(q) \cdot q + p(q)$$

Because demand is downward-sloping, $p'(q) < 0$, selling more units requires lowering the price. If $p'(q) < 0$ (i.e., negative) and quantities are always non-negative, the term $p'(q) \cdot q$ has to be negative. This tells us something slightly different: to sell one more unit, the firm must cut the price on every unit it already sells. This is why $MR < p$ for any monopolist: the revenue gained on the new unit is partly offset by the revenue lost on all existing units.³

Now suppose that inverse demand can be written as $p(q) = 60 - q$. Then $p'(q) = -1$: each additional

³Compare this to a firm in a competitive market, in which every firm is a **price-taker**: because each firm is so small relative to total market supply, its individual output decision has no effect on the market price $\bar{p}$. For that firm, $p'(q) = 0$, so the product rule gives $MR(q) = p'(q) \cdot q + p(q) = 0 + \bar{p} = \bar{p}$. Marginal revenue equals price at every unit. A competitive firm loses no revenue on existing sales when it expands output; a monopolist does.

unit sold requires lowering the price by \$1 across the board. Applying the product rule:

$$MR(q) = (-1)(q) + (60 - q) = 60 - 2q$$

What does this tell us? Marginal revenue starts at the same price intercept as demand (when $q = 0$, $MR = 60 = p$) but falls twice as fast. Here is why: selling one more unit brings in $p = 60 - q$ dollars on the new unit, but forces a \$1 price cut on every unit already sold. These two effects together make MR fall at twice the rate of the demand curve.

Rule 8: Quotient Rule

If $h(x) = \frac{f(x)}{g(x)}$ with $g(x) \neq 0$, then:

$$h'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$$

A memory device: “low d-high minus high d-low, over low squared.”

Example: Differentiate $h(x) = \frac{x^2 + 1}{x - 3}$.

Label the numerator $f(x) = x^2 + 1$ and the denominator $g(x) = x - 3$. Their derivatives:

$$f'(x) = 2x \quad g'(x) = 1$$

Plug into the formula $h'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$:

$$h'(x) = \frac{2x \cdot (x - 3) - (x^2 + 1) \cdot 1}{(x - 3)^2} = \frac{2x^2 - 6x - x^2 - 1}{(x - 3)^2} = \frac{x^2 - 6x - 1}{(x - 3)^2}$$

Alternative: use the product rule instead. Any ratio $f(x)/g(x)$ can be rewritten as $f(x) \cdot [g(x)]^{-1}$ and differentiated without the quotient rule. Rewrite:

$$h(x) = (x^2 + 1)(x - 3)^{-1}$$

Now label $f(x) = x^2 + 1$ and $g(x) = (x - 3)^{-1}$. For $g'(x)$, the exponent -1 comes to the front, the exponent decreases by one, and we multiply by the derivative of the inner function $(x - 3)$, which is 1:

$$g'(x) = -1 \cdot (x - 3)^{-2} \cdot 1 = -(x - 3)^{-2}$$

By the product rule:

$$h'(x) = 2x \cdot (x - 3)^{-1} + (x^2 + 1) \cdot [-(x - 3)^{-2}] = \frac{2x}{x - 3} - \frac{x^2 + 1}{(x - 3)^2}$$

Combining over the common denominator $(x - 3)^2$:

$$h'(x) = \frac{2x(x - 3) - (x^2 + 1)}{(x - 3)^2} = \frac{x^2 - 6x - 1}{(x - 3)^2}$$

Same answer. The quotient rule is a convenience formula for this common pattern.

Application: average cost

A firm's average cost is $AC(q) = C(q)/q$: the total cost of producing q units, spread over those q units. Marginal cost is $MC(q) = C'(q)$: the cost of producing one additional unit. The quotient rule connects the two.

To differentiate $AC(q) = \frac{C(q)}{q}$, identify the pieces:

- $f(q) = C(q)$ (numerator: total cost)
- $g(q) = q$ (denominator: quantity)

Now find the derivative of each:

- $f'(q) = C'(q) = MC(q)$: the derivative of total cost is marginal cost, by definition.
- $g'(q) = 1$: the derivative of q with respect to q is simply 1.

Plugging into the quotient rule formula $\frac{f'(q)g(q) - f(q)g'(q)}{[g(q)]^2}$:

$$AC'(q) = \frac{MC(q) \cdot q - C(q) \cdot 1}{q^2}$$

Now notice that $C(q)/q = AC(q)$. Divide numerator and denominator by q :

$$AC'(q) = \frac{MC(q) \cdot q - C(q)}{q^2} = \frac{1}{q} \left[MC(q) - \frac{C(q)}{q} \right] = \frac{MC(q) - AC(q)}{q}$$

The sign of $AC'(q)$ depends entirely on whether MC is above or below AC :

- If $MC < AC$: the next unit costs less than the current average, so it pulls the average down. $AC'(q) < 0$.
- If $MC = AC$: the next unit costs exactly the current average, leaving it unchanged. $AC'(q) = 0$.
- If $MC > AC$: the next unit costs more than the current average, so it pulls the average up. $AC'(q) > 0$.

A numerical check. Let $C(q) = q^2 + 4$. Then $AC(q) = q + 4/q$ and $MC(q) = 2q$. Applying the formula:

$$AC'(q) = \frac{MC(q) - AC(q)}{q} = \frac{2q - (q + 4/q)}{q} = \frac{q - 4/q}{q} = 1 - \frac{4}{q^2}$$

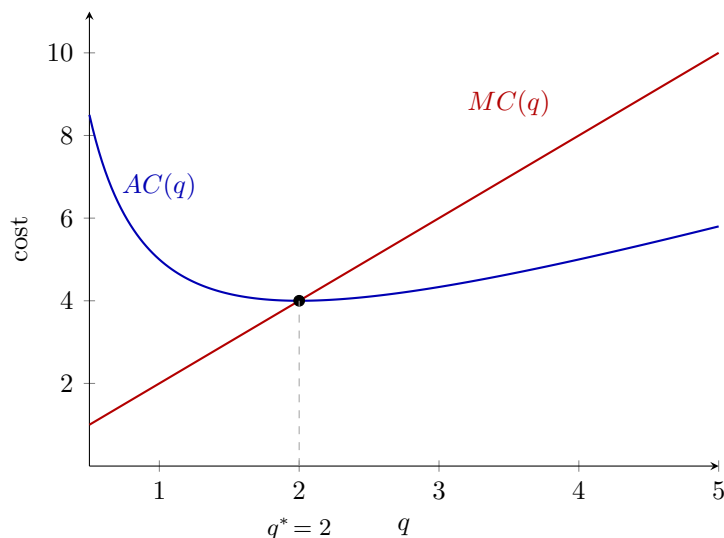


Figure 4: Average and marginal cost: $C(q) = q^2 + 4$.

Notes: $AC(q) = q + 4/q$ (blue) is U-shaped; $MC(q) = 2q$ (red) is increasing. They cross at $q^* = 2$, exactly the minimum of AC . For $q < 2$, $MC < AC$ and average cost is falling; for $q > 2$, $MC > AC$ and average cost is rising.

Setting $AC'(q) = 0$: $q^2 = 4$, so $q^* = 2$. At $q = 2$: $AC(2) = 4$ and $MC(2) = 4$. They are equal at the minimum of AC , as predicted. Figure 4 makes this visible.

This is the mathematical statement of a result you may have seen in microeconomics: the MC curve always passes through the minimum of the AC curve. The quotient rule delivers the proof in a single calculation.

Rule 9: Chain Rule

If $y = f(u)$ and $u = g(x)$, so that $y = f(g(x))$, then:

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx} = f'(g(x)) \cdot g'(x)$$

A change in x changes u (at rate $g'(x)$), and a change in u changes y (at rate $f'(u)$). The total effect on y is the product of these two rates. In economics, this often describes an indirect effect: x shifts some intermediate quantity, which then shifts the outcome.

Example 1: Differentiate $y = (3x^2 + 5)^4$.

Let $u = 3x^2 + 5$ (the “inside”) and $y = u^4$ (the “outside”). Then $du/dx = 6x$ and $dy/du = 4u^3$.

$$\frac{dy}{dx} = 4u^3 \cdot 6x = 4(3x^2 + 5)^3 \cdot 6x = 24x(3x^2 + 5)^3$$

Example 2: Differentiate $y = (3x - x^3)^6$.

$$\frac{dy}{dx} = 6(3x - x^3)^5 \cdot (3 - 3x^2)$$

After a little practice, you can skip writing out u explicitly and work directly.

Application: factor demand

Suppose a firm's production function is $q = L^{1/2}$, where L is employment. Capital is fixed in the short run (a given factory size), so labor is the only variable input. The firm's labor demand depends on the wage rate: $L(w) = 50 - 2w$.

When the wage rate increases, labor becomes more expensive and the firm economizes: it reduces hiring. But fewer workers means less output. The firm does not simply absorb the higher cost; it adjusts. The question is how much output falls per dollar of wage increase.

This is a chain of effects: $w \rightarrow L \rightarrow q$. The chain rule captures both links:

$$\frac{dq}{dw} = \underbrace{\frac{dq}{dL}}_{\text{output response to labor}} \cdot \underbrace{\frac{dL}{dw}}_{\text{labor response to wage}}$$

- $\frac{dq}{dL} = \frac{1}{2}L^{-1/2}$: each additional worker raises output, but with diminishing returns (the more workers already hired, the smaller the marginal contribution of one more).
- $\frac{dL}{dw} = -2$: a \$1 wage increase reduces employment by 2 workers.

Multiplying:

$$\frac{dq}{dw} = \frac{1}{2}L^{-1/2} \cdot (-2) = -\frac{1}{\sqrt{L}}$$

At $L = 25$: $dq/dw = -1/5$. A \$1 wage increase reduces output by 0.2 units.

Could we have plugged in directly? Yes. Substituting $L = 50 - 2w$ into the production function gives $q = (50 - 2w)^{1/2}$. Differentiating (using the chain rule again):

$$\frac{dq}{dw} = \frac{1}{2}(50 - 2w)^{-1/2} \cdot (-2) = -\frac{1}{\sqrt{50 - 2w}} = -\frac{1}{\sqrt{L}}$$

Same answer. Keeping the result in terms of L is a matter of interpretive choice, not a mathematical necessity. The choice really depends on what feature of the model you want to extract.

Why keep it in terms of L ? The expression $-1/\sqrt{L}$ is more informative than $-1/\sqrt{50 - 2w}$. It shows that the wage impact on output, *on the margin*, depends on how many workers the firm currently employs. In this example, a labor-intensive firm (large L) sits on a flatter part of the production curve: each additional worker it has hired is adding less to output than the one before, the marginal workers are basically slack. Diminishing returns are stronger there, so a wage shock reduces output less.⁴ A firm with very few workers is on the steep part of the curve: every worker

⁴It does not automatically follow that a large wage *increase* costs a labor-intensive firm less output overall; that would require summing dq/dw over the whole range the wage moves through, not just evaluating it once. It happens to be true here because $q(w) = \sqrt{50 - 2w}$ gets steeper (in magnitude) as w rises: diminishing returns compound as workers are cut, so the marginal loss only grows as L shrinks. A \$1 raise starting from $L = 50$ costs about 0.14 units of output; the same \$1 raise starting from $L = 2$ costs about 0.71 units.

still on staff is highly productive, so losing even one or two has a large effect on output.

Here is a way to see why. A \$1 wage increase raises the wage bill by \$1 for every worker on payroll: a \$50 increase in total wage costs for the large firm ($L = 50$), but only a \$2 increase for the small firm ($L = 2$). By dollar cost alone, the large firm is hit much harder. But the two firms can respond differently to that cost. The large firm can absorb its (much bigger) wage bill increase by letting go of its least productive workers, the ones diminishing returns had already made nearly redundant, so output barely moves. The small firm has no such workers to spare: both of its employees are essential, so any cut comes out of a worker who really matters. The firm facing the larger dollar cost increase ends up with the smaller output loss, because it can pay for that cost out of workers it was barely using anyway.

This is a case where the algebra reveals a relationship that isn't obvious just from intuition: the firm facing the bigger dollar cost is not the one that suffers the bigger output loss.

Practice Problems

1. Differentiate each of the following functions using the appropriate rules.

(a) $f(x) = 7$

(b) $f(x) = 4x - 9$

(c) $f(x) = 3x^4$

(d) $f(x) = x^5 - 3x^3 + 2x - 8$

(e) $f(x) = \frac{1}{x^2}$ (write as a power function first)

(f) $f(x) = 4\sqrt{x} - \frac{2}{x}$

2. A firm has total cost $C(q) = q^2 - 6q + 30$.

(a) Find marginal cost $MC(q)$.

(b) At what output is marginal cost equal to 0? What does this mean?

(c) Is marginal cost increasing or decreasing for $q > 3$? How can you tell from the derivative?

3. A monopolist faces inverse demand $P(q) = 80 - 2q$.

(a) Write total revenue $TR(q)$ as a function of q only.

(b) Find marginal revenue $MR(q)$.

(c) At what output does $MR = 0$? What is happening to total revenue at that point?

4. A firm's production function is $q = \sqrt{K}$, where K is the number of machines it operates (labor is fixed in the short run, so capital is the only variable input). The firm rents its machines, and its capital stock is $K(r) = 36 - 3r$, where r is the rental rate, in dollars per machine.

- (a) Using the chain rule, find dq/dr as a function of K . Make the decomposition explicit as dq/dK and dK/dr .
- (b) At $r = 9$, how many machines does the firm operate? Evaluate dq/dr at this point, and interpret the number in words.
- (c) Confirm your answers to (a) and (b) by substituting $K(r)$ directly into the production function and differentiating with respect to r .
- (d) Would a capital-intensive firm (large K) or a firm with very little capital (small K) suffer a bigger marginal drop in output from a one-dollar increase in the rental rate? Explain using the idea of diminishing returns. **Hint:** pick two numbers like $K = 3$ and $K = 30$.