

8

CHAPTER 8

Production

How firms turn inputs into output, from a food truck to a tens-of-billions-of-dollars chip foundry.

What this chapter covers



8.2–8.3 Short run

- The production function
- Total, marginal, average product
- Diminishing returns



8.4–8.5 Long run

- Isoquants & substitution
- Returns to scale
- Scale vs. diminishing product



8.6–8.8 Progress

- Kinds of technological progress
- Case: chip foundry (TSMC)
- Empirical: Moore's law

THE BIG IDEA

Production is how firms turn inputs into output, the foundation for the analysis of costs in Chapter 9.

8.1

Introduction

- Why some goods need vast input combinations
- Why output eventually rises more slowly
- How chips became millions of times more powerful

Three questions about production



Scale

Why does a chip factory cost tens of billions while a food truck needs one cook and a griddle?



Crowding

When a restaurant crowds more cooks into one kitchen, why does output eventually rise more slowly?



Progress

How have computer chips become millions of times more powerful within a single lifetime?

WHAT TIES THEM TOGETHER

These questions are all about production, how firms turn inputs into output, and they set up the analysis of costs in Chapter 9.

8.2

The Production Function and the Short Run

- Inputs: labor and capital
- The short run vs. the long run

8.2 PRODUCTION FUNCTION

Inputs in, output out

■ Production function

Describes the maximum output a firm can produce from each combination of inputs.

■ Two broad kinds of input

Labor, the workers. Capital, the machines, buildings, and equipment.

8.2 PRODUCTION FUNCTION

Two time horizons



Short run

At least one input is fixed, typically capital, since a firm cannot build a new factory overnight.



Long run

Every input can be varied: the firm can build plants, install machines, and hire freely.

THE KEY DISTINCTION

The short run is about working an existing factory harder; the long run is about choosing the factory itself. A restaurant can add cooks tonight but cannot enlarge its kitchen until next year.

8.3

Total, Marginal, and Average Product

- The production table
- Marginal & average product
- Diminishing marginal product
- The shape of the curves

8.3 TOTAL, MARGINAL, AVERAGE PRODUCT

Adding cooks to a fixed kitchen

Workers	Total product	Marginal product	Average product
0	0		
1	10	10	10.0
2	24	14	12.0
3	39	15	13.0
4	50	11	12.5
5	58	8	11.6
6	63	5	10.5

Table 8.1 Total, marginal, and average product as a kitchen adds cooks.

WHAT THE NUMBERS SHOW

In the short run, with capital fixed, the firm raises output, the total product, by adding labor.

Marginal product first rises (10, 14, 15) as early cooks specialize, then falls (11, 8, 5).

8.3 TOTAL, MARGINAL, AVERAGE PRODUCT

Two derived numbers do most of the work

■ Marginal product of labor

The extra output from one more worker, the third cook adds 15 meals.

■ Average product

Total output divided by the number of workers, with three cooks producing 39 meals, the average is 13 each.

The law of diminishing marginal product

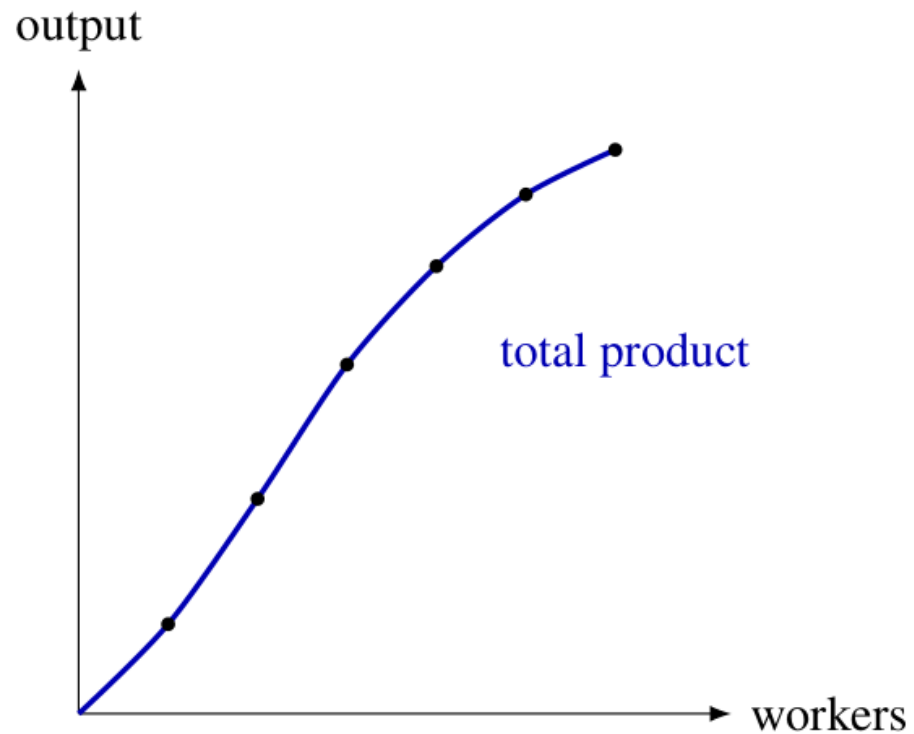
THE LAW

As more of one input is added to fixed amounts of others, the extra output from each additional unit eventually shrinks.

- With the kitchen fixed, the fourth, fifth, and sixth cooks crowd the same stoves and counters.
- So each adds less output than the one before, marginal product falls from 15 to 11 to 8 to 5.

8.3 TOTAL, MARGINAL, AVERAGE PRODUCT

The total-product curve

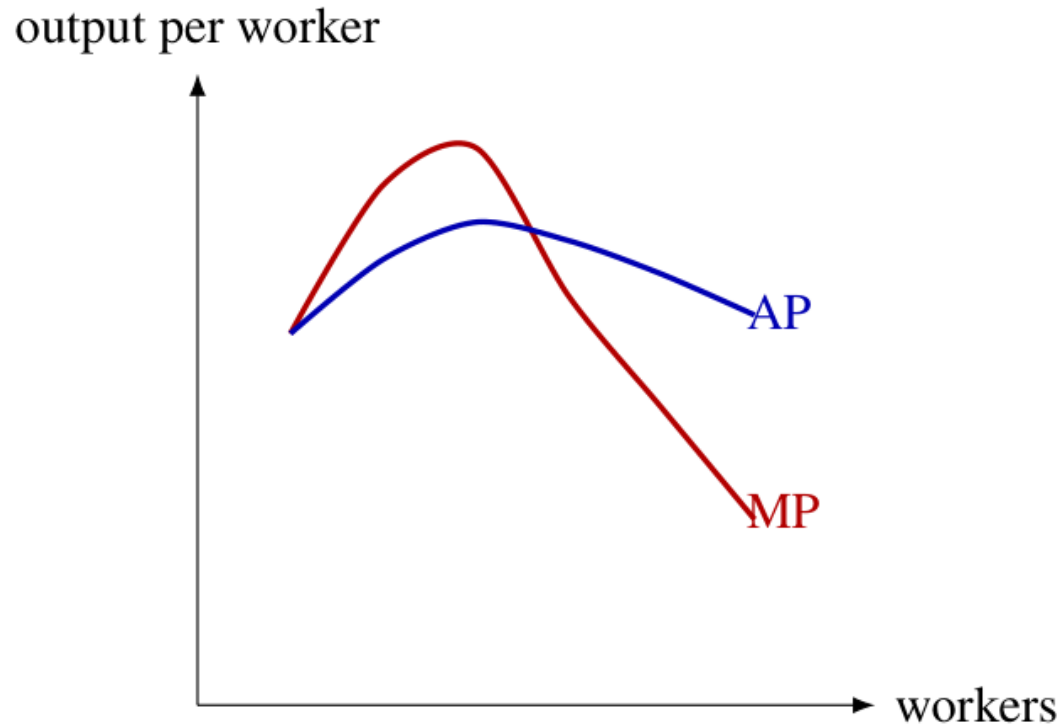


- Early on the curve is steep, each cook adds more than the last.
- Beyond some point it flattens: still rising, but ever more slowly.

Figure 8.1 Total product rises quickly, then more slowly as diminishing marginal product sets in.

8.3 TOTAL, MARGINAL, AVERAGE PRODUCT

Marginal and average product



- Marginal product rises, then falls.
- It cuts average product at the average's peak.
- When the next worker exceeds the current average, the average rises; when worse, it falls.

Figure 8.2 Marginal product rises then falls, cutting the average-product curve at its maximum.

8.3.1 DIMINISHING RETURNS & THE CURVES

Why the curve steepens, then flattens



Early: rising

With a nearly empty kitchen, each extra cook lets the others specialize, one preps, one cooks, one plates, so output climbs faster.



Later: falling

The kitchen fills; each added cook has less equipment and space, so marginal product falls and total product rises ever more slowly.

THE EXTREME CASE

A kitchen so crowded that workers trip over one another could even push marginal product to zero or below, an extra cook then **LOWERS** total output.

8.3.1 DIMINISHING RETURNS & THE CURVES

Self-check: reading Table 8.1

SELF-CHECK

The third cook raises output from 24 to 39 meals. What is the marginal product of the third cook, and the average product with three cooks?

Marginal product = $39 - 24 = 15$ meals. Average product = $39 / 3 = 13$ meals per cook.

8.4

The Long Run: Substituting Between Inputs

- The same output from different input mixes
- Isoquants
- The marginal rate of technical substitution

Different mixes, same output

■ Isoquant (“equal quantity”)

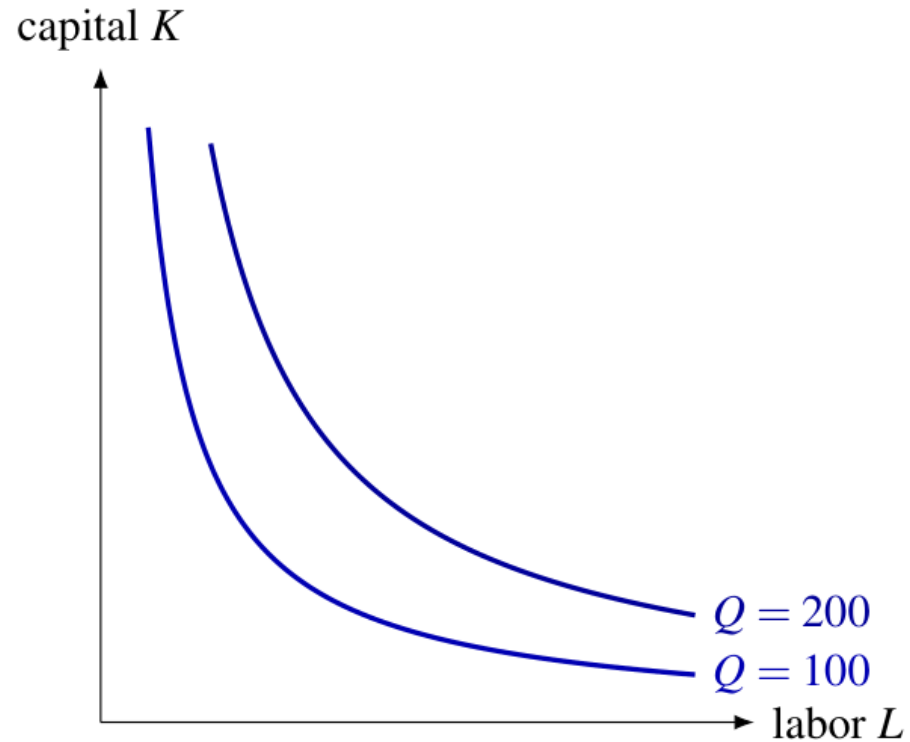
Shows all the input combinations that produce a given level of output.

THE LONG-RUN CHOICE

In the long run a firm can vary both labor and capital, and the same output can usually be produced with different mixes, a car can be assembled by many workers with simple tools, or a few workers with many robots.

8.4 THE LONG RUN

Reading an isoquant map



- Higher isoquants mean more output.
- They slope downward and bow toward the origin, like the indifference curves of Chapter 5.

Figure 8.3 Each isoquant shows the capital–labor combinations that produce a given output.

The marginal rate of technical substitution

■ Marginal rate of technical substitution

The rate at which a firm can trade one input for another while holding output constant.

WHY IT DIMINISHES

It diminishes: a firm with many machines and few workers can give up several machines for one more worker, but as workers grow plentiful each substitutes for fewer and fewer machines. This is why factories settle on a balanced mix rather than relying entirely on one input.

8.5

Returns to Scale

- Scaling ALL inputs up together
- Increasing, constant, decreasing
- Scale vs. diminishing marginal product

8.5 RETURNS TO SCALE

What if a firm doubles every input?



Increasing

Doubling all inputs MORE than doubles output. Common at low volumes, larger operations allow specialization and bigger, more efficient machines.



Constant

Doubling all inputs exactly doubles output, for example, building a second identical factory beside the first.



Decreasing

Doubling all inputs LESS than doubles output, often because very large organizations become harder to coordinate and manage.

Returns to scale ask what happens when ALL inputs scale up together, not just one.

8.5 RETURNS TO SCALE

Why it matters

LOOKING AHEAD

Returns to scale shape the cost curves of Chapter 9, and help explain why some industries are dominated by a few enormous firms while others support many small ones.

8.5.1 TWO IDEAS, KEPT APART

Returns to scale vs. diminishing marginal product

	Diminishing marginal product	Returns to scale
Inputs changed	one input; others fixed	all inputs together
Example	more cooks, same kitchen	bigger kitchen AND more cooks
When	always sets in eventually	may increase, be constant, or decrease
Horizon	the short run	the long run

The two are entirely compatible, do not confuse them.

8.5.1 TWO IDEAS, KEPT APART

The two are compatible

NO CONTRADICTION

A firm can have increasing returns to scale (doubling everything more than doubles output) and still face diminishing marginal product (adding cooks alone, kitchen fixed, runs into crowding).

- Mixing them up leads to the false belief that diminishing marginal product rules out economies of scale.
- It does not, the giant, low-cost factories of Chapter 9 will show why.

8.6

Technological Progress

- Shifting the production function outward
- Three kinds of progress
- Progress and workers

8.6 TECHNOLOGICAL PROGRESS

Shifting the production function

■ Technological progress

Over time, the firm produces more output from the same inputs, as better methods, machines, and knowledge arrive.

WHY IT MATTERS

Technological progress is the main reason living standards rise over the long run, and in no industry has it been faster or more measurable than in semiconductors.

8.6.1 KINDS OF TECHNOLOGICAL PROGRESS

Three kinds of progress



Neutral

Raises output from labor and capital in equal proportion, a better recipe that lets the same cooks and ovens make more meals.



Labor-saving

Raises capital's productivity relative to labor, so firms substitute machines for workers, the industrial robots doing work once done by hand.



Capital-saving

Does the reverse, making labor relatively more productive.

The differences matter for workers, and for the automation debate.

Progress and workers

THE TENSION

Most automation anxiety concerns labor-saving progress, which can displace workers from particular jobs even as it makes the whole economy more productive.

- Economics offers no guarantee that everyone gains from technological change in the short run.
- What it shows: rising productivity expands the total that CAN be shared, how it is shared is a question for policy and politics.

8.7

Case Study: Inside a Chip Foundry

TSMC, the economics of production, made vivid.

What is a foundry?

■ Foundry

A firm that manufactures chips designed by other companies, from smartphone processors to the graphics chips behind artificial intelligence.

~60%

of the global foundry market held by TSMC in 2023

Most

of the world's most advanced chips, produced by TSMC

Capital-intensive, and built on scale



Extreme capital intensity

A single leading-edge fab, a “fab”, costs tens of billions, far more than the labor that runs it. An extreme point of the isoquant (Section 8.4).



Powerful increasing returns

The enormous fixed cost of designing and building a fab is the same whether it makes a few chips or millions, so larger volumes sharply lower cost per chip.

THE UPSHOT

This is why a handful of giant firms, not many small ones, make the world's advanced chips, and why TSMC's scale is so hard for rivals to match.

8.8

Empirical Evidence: Moore's Law

The most famous quantitative record of technological progress.

8.8 EMPIRICAL EVIDENCE

Moore's law

■ Moore's law

In 1965 Gordon Moore noticed transistors per chip had been doubling about every year; in 1975 he revised the pace to a doubling roughly every two years.

~2 yrs

a doubling of transistors per
chip

~40%

growth per year, exponential

8.8 EMPIRICAL EVIDENCE

Half a century of doubling

- The prediction held for about half a century, turning chips with thousands of transistors into chips with tens of billions.
- Each doubling was the production function of Section 8.6 shifting outward: the same silicon wafer, processed with better methods, yielded vastly more computing power.

WHAT IT REALLY IS

Moore's law is not a law of nature but the cumulative result of sustained investment and engineering. The pace has recently slowed as transistors approach the size of atoms, yet it remains the clearest evidence of how powerful, and how rapid, technological progress can be.

Key takeaways

1

Production function

Maximum output from each input combination, labor and capital.

2

Short vs. long run

Short run: at least one input fixed. Long run: every input varies.

3

Marginal & average product

Diminishing marginal product always sets in as one input is added to fixed others.

4

Isoquants

Show input mixes for a given output; the MRTS diminishes, so firms balance inputs.

5

Returns to scale

Scaling ALL inputs may give increasing, constant, or decreasing returns, different from diminishing product.

6

Technological progress

Shifts the function outward; Moore's law is its most famous record.