Borjas Textbook Notes

Dylan Baker

September 10, 2025

Contents

1 Chapter 1: Introduction

1.1 The Actors in the Labor Market

This book will consider three actors within the labor market: (1) workers, (2) firms, and (3) government. The decisions of workers will be based on a desire to optimize what Borjas calls their well-being. These decisions across workers generate the labor supply curve. Firms have the goal of maximizing profits. The firm's demand for labor is a "derived demand" in the sense that it is derived from consumer's demand for the firm's output. Equilibrium is attained when supply equals demand in a free market economy. The government's motives are left more opaque.

1.2 Why Do We Need a Theory?

Writing out supply and demand curves reflects the construction of a model, which makes predictions about what will transpire if certain conditions change. The model is simple but is useful for organizing our thoughts about the labor market and provides a solid foundation upon which to build more complex infrastructure.

The predictions of the supply and demand model is an example of positive economics. Positive economics is concerned with "What is?" questions, i.e., questions about how the world actually works. This is in contrast to normative economics, which is concerned with questions of "What ought to be?"

2 Chapter 2: Labor Supply

The aggregate labor supply is the sum of the individual labor supply decisions of all prospective workers in the economy. This chapter will be focused on fleshing out the framework that economists use to think about labor supply decisions.

2.1 Terms

- LF: The size of the labor force
- E: The number of employed individuals
- U: The number of unemployed individuals
- P: The size of the population
- C: The consumption of goods
- L: The consumption of leisure
- $U = f(C, L)$: The utility function
- MU_L: The marginal utility of leisure
- MU_C: The marginal utility of consumption
- V: Nonlabor income
- h: Hours worked
- w: The wage rate
- T: Total time available in, say, a week
- σ: The labor supply elasticity

2.2 Measuring the Labor Force

The CPS classified individuals over the age of 16 into three categories:

1. Employed: Somebody working at least 1 hour of paid labor or 15 hours of unpaid labor
2. Unemployed: Somebody temporarily laid off from a job or actively looking for work
3. Out of the labor force: Everyone else

Definition D.1: Labor Force

The labor force (LF) is everyone who is employed (E) or unemployed (U).

$$LF = E + U$$

Definition D.2: Labor Force Participation Rate

The Labor Force Participation Rate is the fraction of the population (P) that is in the labor force.

$$\text{Labor force participation rate } = \frac{LF}{P}$$

Definition D.3: Employment Rate

The fraction of the population that is employed.

$$\text{Employment rate } = \frac{E}{P}$$

Definition D.4: Unemployment Rate

The fraction of the population that is unemployed.

$$\text{Unemployment rate } = \frac{U}{LF}$$

Notice that the number of unemployed people is calculated as a fraction of the labor force, not as a fraction of the population. Thus, one way for the unemployment rate to go down is for people to stop looking for work entirely.

2.3 Basic Facts about Labor Supply

Below are just a few facts of key labor supply trends over the last century. Figure 1 shows the labor force participation rate for men and women over the last century. Figure 2 shows the average weekly hours worked for all workers over the last century.

TABLE 2-1 Labor Force Participation Rates of Men, 1900–2010

Sources: U.S. Bureau of the Census, *Historical Statistics of the United States, Colonial Years to 1970,* Washington, DC: Government Printing Office, 1975; U.S. Bureau of the Census, *Statistical Abstract of the United States,* Washington, DC: Government Printing Office, various issues.

Year	All Men	Men Aged 25–44	Men Aged 45–64	Men Aged over 65
1900	80.0	94.7	90.3	63.1
1920	78.2	95.6	90.7	55.6
1930	76.2	95.8	91.0	54.0
1940	79.0	94.9	88.7	41.8
1950	86.8	97.1	92.0	45.8
1960	84.0	97.7	92.0	33.1
1970	80.6	96.8	89.3	26.8
1980	77.4	93.0	80.8	19.0
1990	76.4	93.3	79.8	16.3
2000	74.8	93.1	78.3	17.5
2010	71.2	90.6	78.4	22.1

TABLE 2-2 Labor Force Participation Rates of Women, 1900–2010

Sources: U.S. Bureau of the Census, *Historical Statistics of the United States, Colonial Years to 1970,* Washington, DC: Government Printing Office, 1975, p. 133; and U.S. Department of Commerce, *Statistical Abstract of the United States, 2011,* Washington, DC: Government Printing Office, 2011, Table 596.

Year	All Women	Single Women	Married Women	Widowed, Divorced, or Separated
1900	20.6	43.5	5.6	32.5
1910	25.4	51.1	10.7	34.1
1930	24.8	50.5	11.7	34.4
1940	25.8	45.5	15.6	30.2
1950	29.0	46.3	23.0	32.7
1960	34.5	42.9	31.7	36.1
1970	41.6	50.9	40.2	36.8
1980	51.5	64.4	49.9	43.6
1990	57.5	66.7	58.4	47.2
2000	59.9	68.9	61.1	49.0
2010	58.6	63.3	61.0	48.8

Figure 1: Labor Force Participation Rate for Men and Women

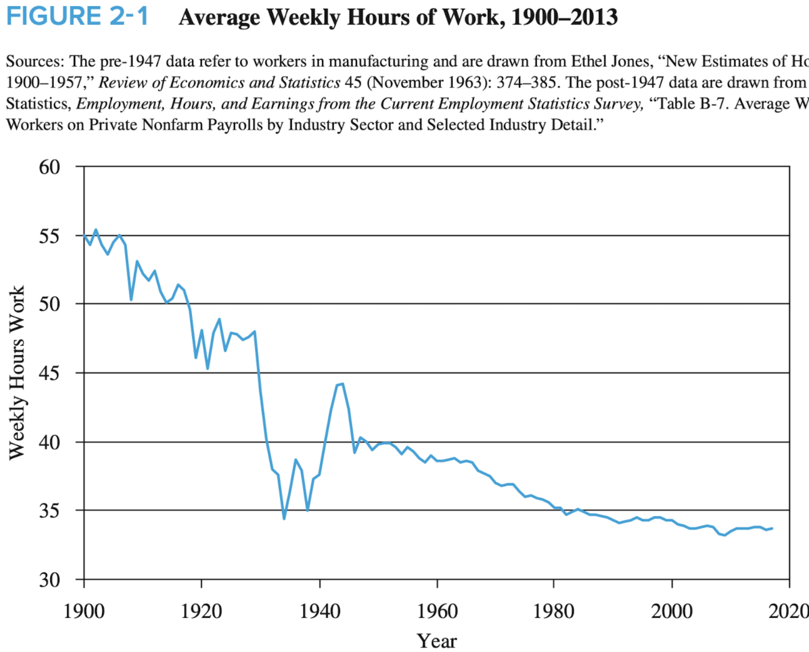

Figure 2: Average Weekly Hours Worked

2.4 The Worker's Preferences

The framework that economist's typically use to think about labor supply behavior is the "neoclassical model of labor-leisure choice." This model is used to think about what's underpinning a worker's decision regarding whether and how much to work.

2.4.1 Indifference Curves

Under this model, we think of an agent as deriving utility from both consumption (C) and leisure (L). We express the utility function as $U = f(C, L)$, where higher U indicates higher utility and is preferred by the agent.

We use indifference curves to represent the combinations of C and L that yield the same level of utility.

See Figure 3 for an example of indifference curves.

FIGURE 2-2 Indifference Curves
Points X and Y lie on the same indifference curve and yield the same utility (25,000 utils); point Z lies on a higher indifference curve and yields more utility.

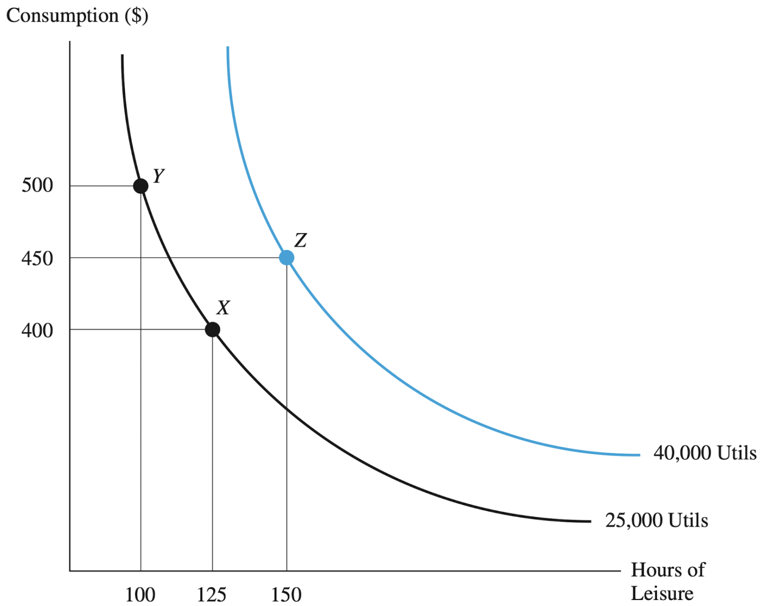

Figure 3: Indifference Curves

Indifference curves can be quite flexible in some aspects of their shape, but there are four properties that we generally impose:

1. Indifference curves are downward sloping: Otherwise, you could get more of both C and L without getting any additional utility.
2. Higher indifference curves indicate higher levels of utility
3. Indifference curves don't intersect: Otherwise, one point would yield two different levels of utility.
4. Indifference curves are convex to the origin

2.4.2 The Slope of an Indifference Curve

Marginal utility from an additional unit of consumption is denoted MU_C. Similarly, the marginal utility from an additional unit of leisure is denoted MU_L.

The slope of an indifference curve is given by

$$\frac{\Delta C}{\Delta L} = -\frac{MU_L}{MU_C}$$

This is logical, since the claim of indifference mandates that the utility change is equivalent as you move along the curve, so it must be that

$$\left| \underbrace{\Delta C}_{\substack{\text{reduction in}\\ \text{consumption}}} \cdot \underbrace{MU_C}_{\substack{\text{marginal}\\ \text{utility of}\\ \text{consumption}}} \right| = \left| \underbrace{\Delta L}_{\substack{\text{increase in}\\ \text{leisure}}} \cdot \underbrace{MU_L}_{\substack{\text{marginal}\\ \text{utility of}\\ \text{leisure}}} \right|$$

That is, you're essentializing making the changes equivalent after scaling by their impact on utility.

Definition D.5: Marginal Rate of Substitution (MRS) in Consumption

The Marginal Rate of Substitution (MRS) in consumption is the absolute value of the slope of the indifference curve.

That is,

$$MRS_{CL} = \frac{MU_L}{MU_C} = \left|\frac{\Delta C}{\Delta L}\right|$$

2.5 The Budget Constraint

Let

- V: Nonlabor income
- h: Hours worked
- w: The wage rate
- L: Hours of leisure

- T: Total hours available in, say, a week

$$T = L + h$$

An individual's budget constraint can then be written in various ways:

$$\begin{aligned} C =& wh + V \\ =& w(T - L) + V \\ =& (wT + V) - wL \end{aligned}$$

The budget constraint then describes the boundary of the worker's opportunity set. This last expression is helpful towards this end, because it follows the familiar $y = mx + b$ format, so we can use it to graph the budget constraint with $(wT + V)$ as the intercept and $-w$ as the slope as leisure increases.

See Figure 4 for an example.

FIGURE 2-5 The Budget Line Is the Boundary of the Worker's Opportunity Set
Point *E* is the endowment point, telling the person how much she can consume if she does not work at all. The worker moves up the budget line as she trades an hour of leisure for consumption of goods. The absolute value of the slope of the budget line is the wage rate.

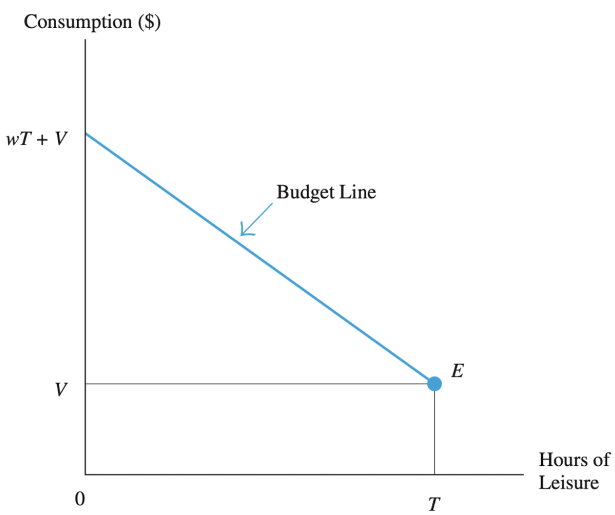

Figure 4: Worker's Opportunity Set

2.6 The Hours of Work Decision

Supposing there is an interior solution, i.e., the worker chooses to work some positive amount of hours but not all available hours, the optimal choice of C and L occurs where the budget constraint is tangent to an indifference curve. See Figure 5 for an example.

FIGURE 2-6 **Interior Solution to the Labor–Leisure Decision**
A utility-maximizing worker chooses the consumption–leisure bundle at point P, where the indifference curve is tangent to the budget line.

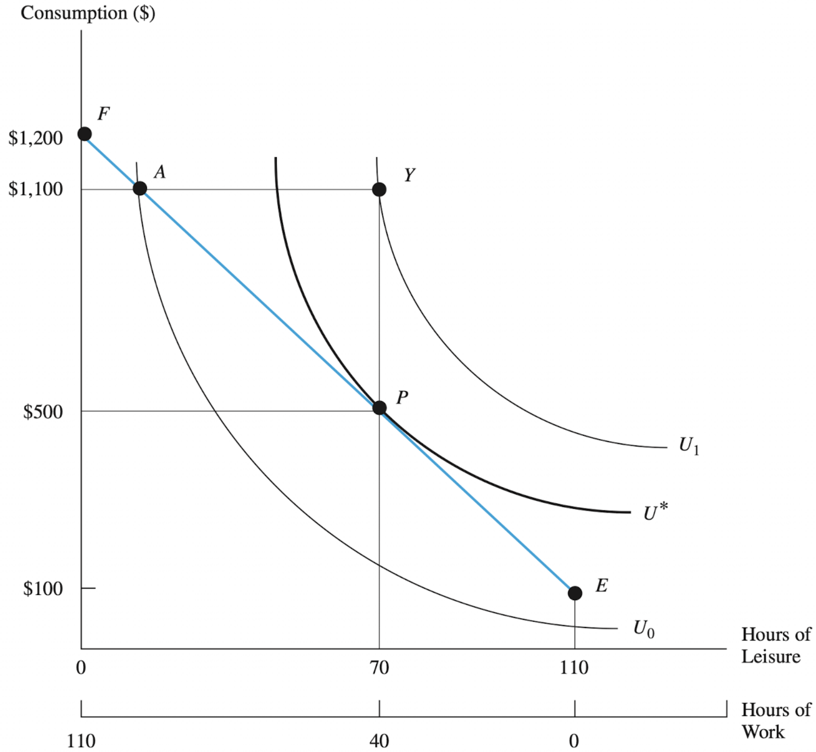

Figure 5: Tangency Condition for Optimal Choice

At the optimal point, denoted by P in Figure 5, the slope of the indifference curve equals the slope of the budget constraint. That is,

$$\frac{MU_L}{MU_C} = w$$

This is easier to interpret, in my opinion, if we re-write it as:

$$MU_L = wMU_C$$

That is, we are indifferent between the marginal utility from an additional unit of leisure (MU_L) and the marginal utility from consumption (MU_C) multiplied by the amount of consumption that you could get from working the extra hour (w).[1] Thus, the last dollar spent on consumption yields the same marginal utility as the last dollar spent on leisure.

[1] I think this is based on a normalization, so that the price of one unit of the consumption good is 1.

2.6.1 What Happens to Hours of Work When Nonlabor Income Changes

I won't spend long on this discussion, but see Figure 6 for a graphical depiction considering what happens to consumption and leisure when nonlabor income (V) increases and leisure is a normal or inferior good. Essentially this changes the frontier of the worker's opportunity set by shifting the intercept up without changing the slope. Thus, it's a pure income effect, with no substitution effect. If leisure is a normal good, which the author suggests is probably the case, then this should lead to a decrease in hours worked.

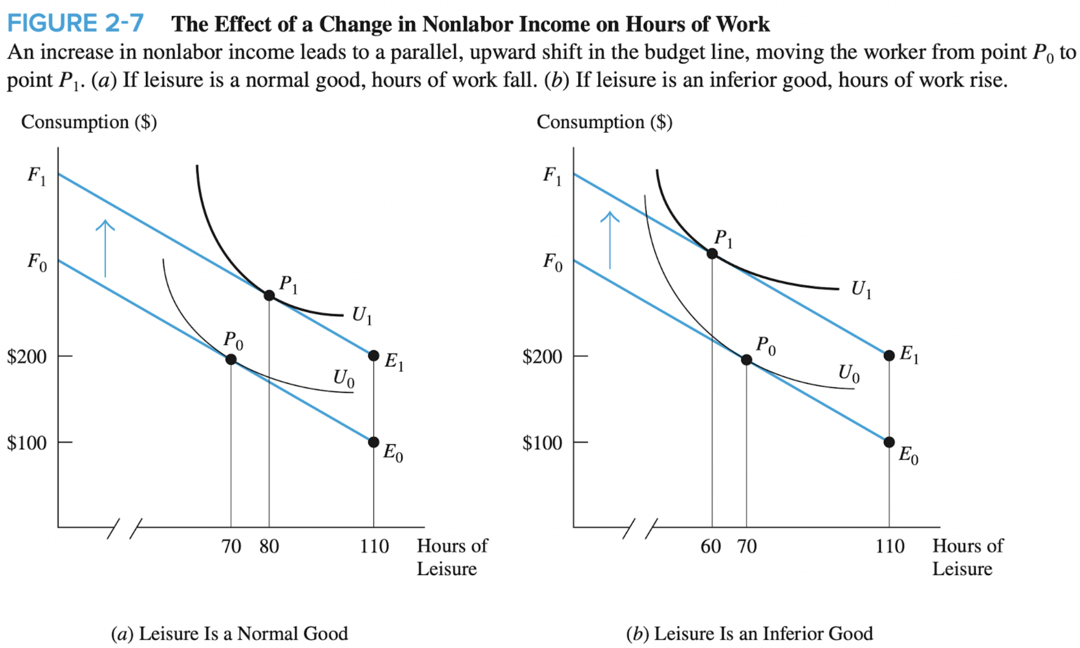

Figure 6: Effect of Nonlabor Income Shock on Hours of Work

2.6.2 What Happens to Hours of Work When the Wage Changes?

If there is a change to the worker's wage, then the effect on hours worked is ambiguous, as we must contend with both income and substitution effects. See Figure 7 for a graphic depiction of what happens to the worker's opportunity set when the wage increases, as well as an example of indifference curves that could generate an increase or decrease in hours worked.

FIGURE 2-8 The Effect of a Change in the Wage Rate on Hours of Work
A change in the wage rate rotates the budget line around the endowment point E. A wage increase moves the worker from point P to point R, and can either decrease or increase hours of work.

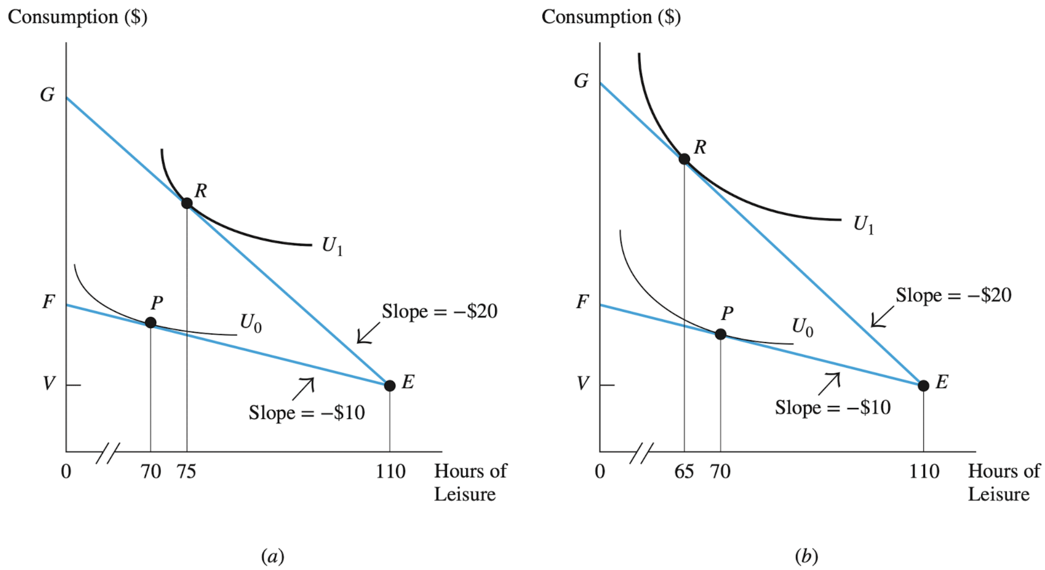

Figure 7: Effect of Wage Change on Hours of Work

Figure 8 decomposes the effect of the wage change into the income and substitution effects. In this decomposition, the pre-wage-change level of consumption and leisure is given by P. After the wage change, the move from P to Q is the income effect, in which the worker works fewer hours due to the increase in income. We visualize this with a parallel shift in the budget constraint,[2] as we did in Figure 6 when the worker experienced a non-wage income shock. The income effect portion is thus characterized by a change in the intercept of the budget constraint line, but not the slope.

The substitution effect is then the move from Q to R, in which the worker works more hours due to the increase in the wage. In contrast to when the worker experienced a non-wage income shock, leisure has now become more expensive, in the sense that the opportunity cost of leisure has increased. The substitution effect can be thought of as what happens to the worker's choice of C and L as the wage changes but the worker's level of utility is held constant, in this case at the level associated with U_1.

As demonstrated in Figure 8, the overall effect of the wage change on hours worked is ambiguous, as it depends on the relative magnitudes of the income and substitution effects.

[2]When I say budget constraint here, I guess I really mean the opportunity set frontier.

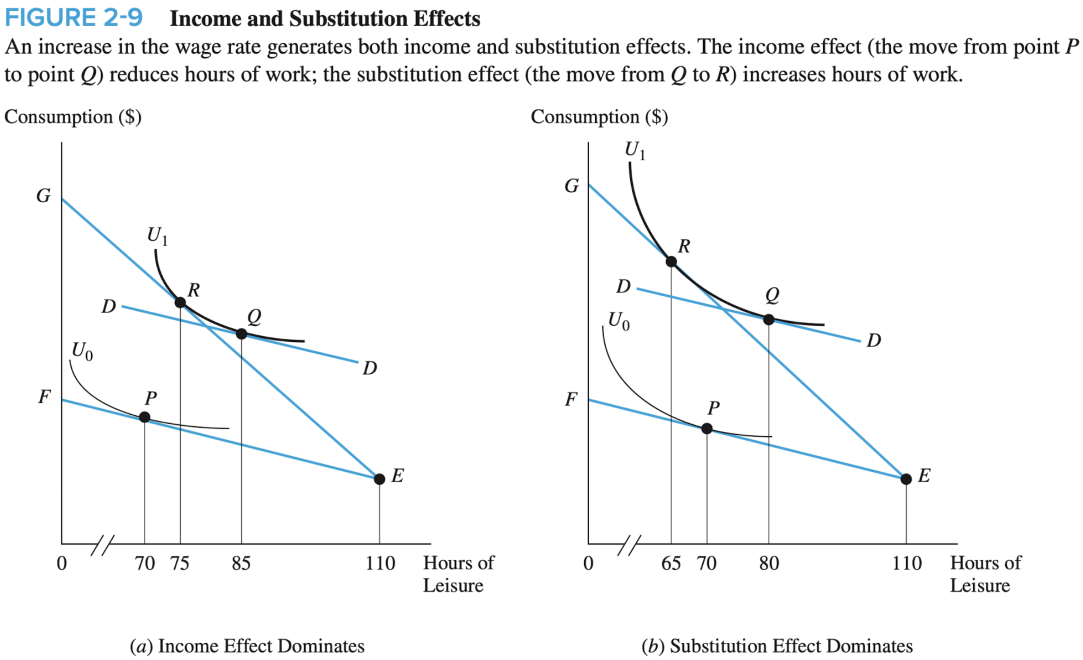

Figure 8: Decomposition of Wage Change Effects

2.7 To Work or Not to Work?

A worker may choose not to work at all. Look at Figure 9 for an example. At wage w_{low}, the worker chooses not to work. This is because the utility curve that they're on at their initial endowment point E (U_0) is higher than any utility curve that they could reach along the line GE characterizing their opportunity set under wage w_{low}. However, if the wage increases to w_{high}, then, they could reach point Y on indifference curve U_H, which is higher than U_0. Thus, they work at wage w_{high}. In fact, we can pinpoint the exact point at which the worker is indifferent between working and not working. This occurs at the wage w^*, which is the absolute value of the slope of the line tangent to the indifference curve U_0 at point E. We refer to this as the reservation wage.

Definition D.6: Reservation Wage

The reservation wage is the wage at which the worker is indifferent between working and not working.

Note that the probability of working is increasing in the wage – only the substitution effect applies, since the income effect of a wage increase doesn't kick in if the person wasn't working to begin with.

FIGURE 2-10 **The Reservation Wage**
If the person chooses not to work, she can remain at the endowment point E and get U_0 units of utility. At a low wage (w_{low}), the person is better off not working. At a high wage (w_{high}), she is better off working. The reservation wage w^* is given by the slope of the indifference curve at the endowment point.

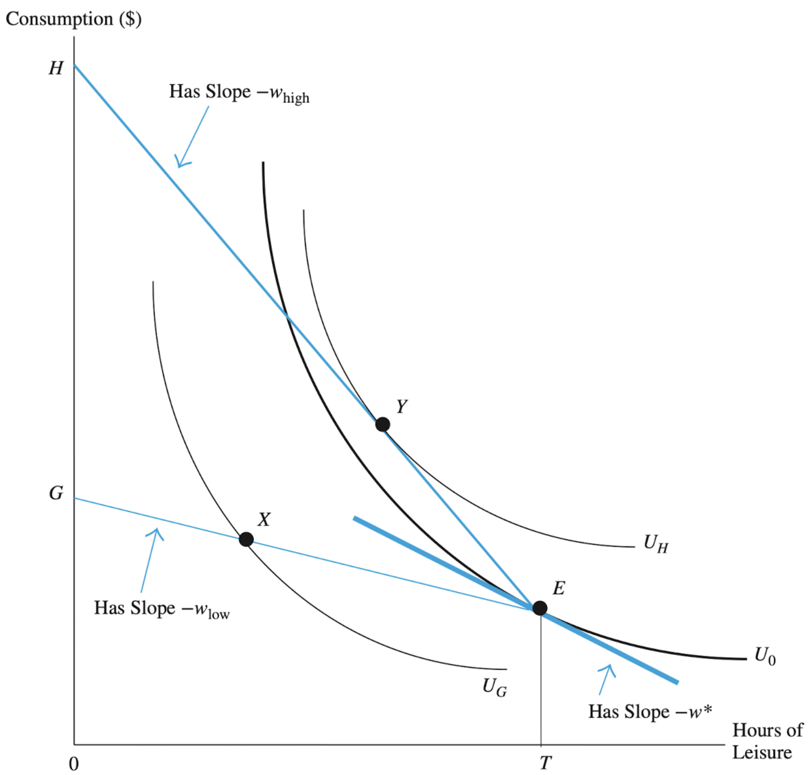

Figure 9: Decision to Work

2.8 The Labor Supply Curve

There is a tight relationship between the utility maximization problem that we've been discussing and the worker's labor supply curve. Indeed, the latter can be derived from the former, as demonstrated in Figure 10. At each wage, the worker chooses a level of hours to spend on leisure (L), the hours worked is then $h = T - L$, where T is the total hours available. Thus, for each wage, we have the worker's corresponding number of hours of labor supplied, which is what we need to plot the labor supply curve. In the example we consider in Figure 10, the substitution effect dominates up until the wage hits 20, after which point, the income effect dominates and the number of hours worked declines.

FIGURE 2-11 Deriving a Labor Supply Curve for a Worker
The labor supply curve traces out the relationship between the wage rate and hours of work. At wages below the reservation wage ($10), the person does not work. At wages higher than $10, the person enters the labor market. The upward-sloping segment of the labor supply curve implies that substitution effects are stronger initially; the backward-bending segment implies that income effects may dominate eventually.

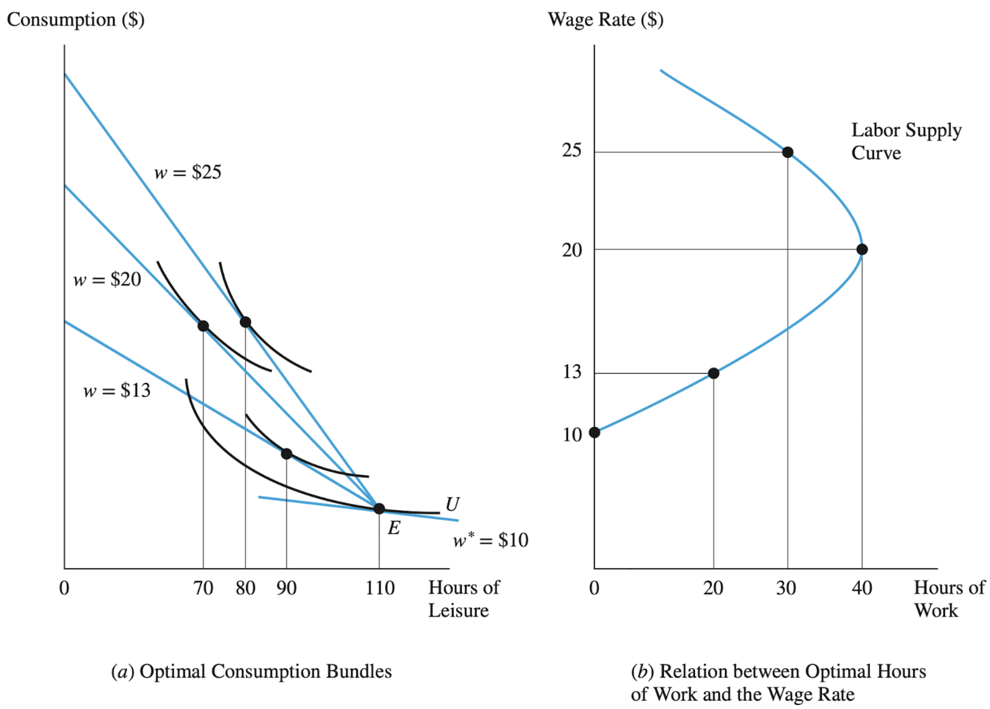

Figure 10: Derivation of the Labor Supply Curve

The aggregate labor supply is the sum of the individual labor supply decisions. See Figure 11 for an example with two workers.

FIGURE 2-12 Derivation of the Market Labor Supply Curve
The market labor supply curve "adds up" the supply curves of individual workers. When the wage is below $15, no one works. At a wage of $15, Alice enters the labor market. If the wage rises above $20, Brenda also enters the market.

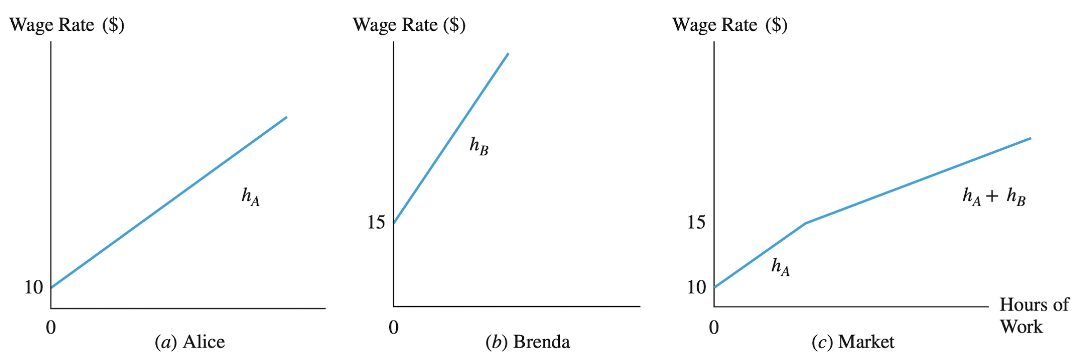

Figure 11: Aggregate Labor Supply

We then consider the notion of labor supply elasticity. Labor supply elasticity tells us what percent change in hours worked corresponds to a percent change in the wage.

Definition D.7: Labor Supply Elasticity

The labor supply elasticity is a measure of how responsive hours worked (h) is to changes in the wage (w). It is given by the expression for σ below.

$$\sigma = \underbrace{\left(\frac{\Delta h}{h}\right)}_{\substack{\text{Percent}\\\text{change in}\\\text{hours}\\\text{worked}}} \Big/ \underbrace{\left(\frac{\Delta w}{w}\right)}_{\substack{\text{Percent}\\\text{change in}\\\text{wage}}} = \frac{\Delta h}{\Delta w} \cdot \frac{w}{h}$$

Definition D.8: Elastic and Inelastic

If $\sigma > 1$, then labor supply is elastic. If $\sigma < 1$, then labor supply is inelastic. If $\sigma = 1$, then labor supply has unit elasticity.

2.9 Estimates of the Labor Supply Elasticity

Some evidence suggests that the labor supply elasticity for men is around -0.1. That is, the income effect is dominating for men, which could explain the decline in hours worked among men over the last century, as real wages have increased. However, there are meaningful challenges in estimating the labor supply elasticity, e.g., measurement error.

2.10 Household Production

The neoclassical model of labor-leisure choice is built on the idea that we can either engage in leisure or wage labor. However, we often engage in home production, in which we produce goods and services for our own consumption. We now incorporate this into a model of labor supply that considers a household (in this case, 2 workers), rather than just an individual.

2.10.1 The Household Production Function

Definition D.9: The Household Production Function

The household production function describes the level of household output producible by the household for a given amount of time.

See Figure 12 for an example. In this example, we consider a couple, Jack and Jill. Panels (a) and (b) show the opportunity sets for Jack and Jill individually, respectively. Panel (c) shows the household production function, which is derived from the individual opportunity sets. In this scenario, we suppose that Jack can earn more in the market and produce less in the household per hour. Thus, the slope of the household opportunity frontier varies as it switches from the marginal productivity coming from Jack to Jill or vice versa.

One of the key implications of this model is that each member of the household should be prioritized for the activity in which they are comparatively better. For example, if the couple needs to move from a world in which they engage exclusively in home production to one in which some hours of work are allocated to the market, this model would prioritize Jack working in the market, since he has the higher wage and lower household productivity.

FIGURE 2-13 **Budget Lines and Opportunity Frontier of Married Couple**
At point *E*, Jack and Jill allocate all their time to the household sector. If they wish to buy market goods, Jack gets a job because he has a relatively higher wage, generating segment *FE* of the frontier. After he allocates all his time to the labor market, Jill then gets a job, generating segment *GF* of the frontier.

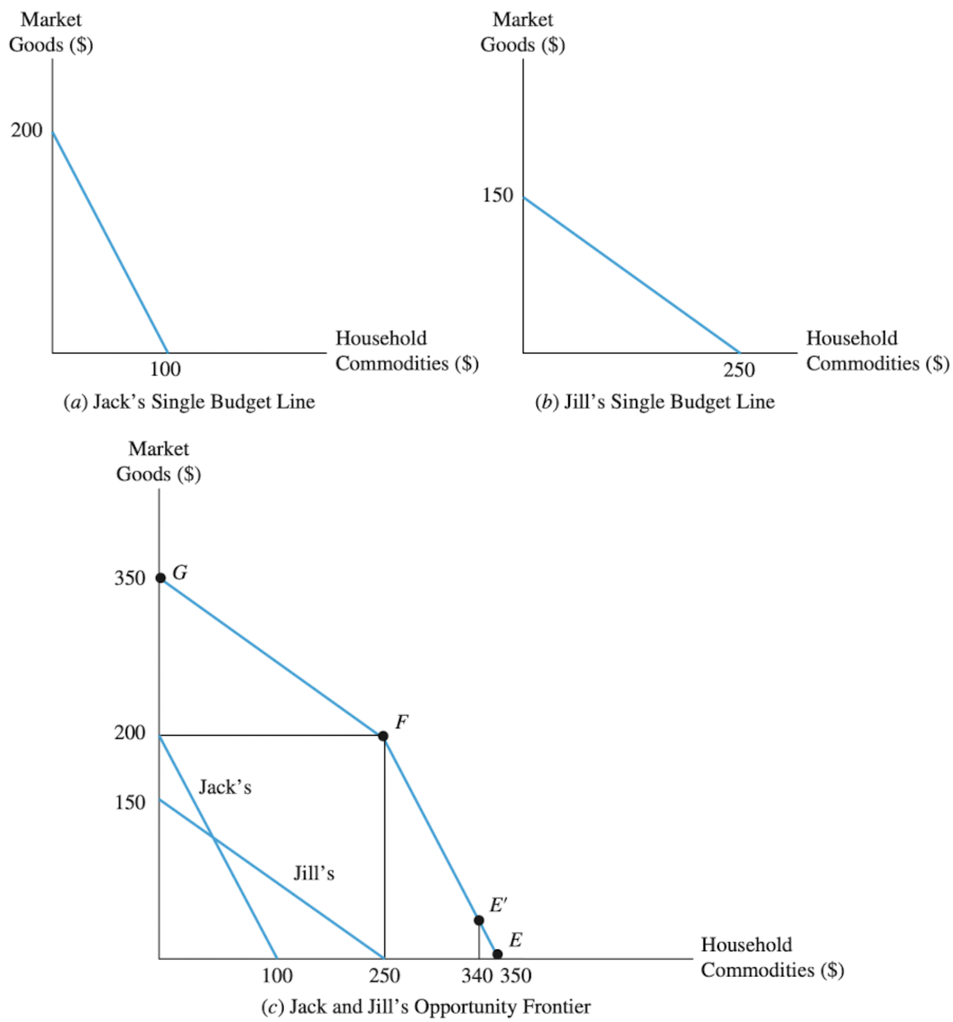

Figure 12: Household Production Function

Figure 13 shows various possible allocations of Jack and Jill's time between market work and home production that might be justified under different utility functions.

FIGURE 2-14 The Division of Labor in the Household

The indifference curve U is tangent to the opportunity frontier at point P. (a) Jill specializes in the household sector and Jack divides his time between the labor market and the household. (b) Jack specializes in the labor market and Jill divides her time between the two sectors. (c) Jack specializes in the labor market and Jill specializes in the household sector.

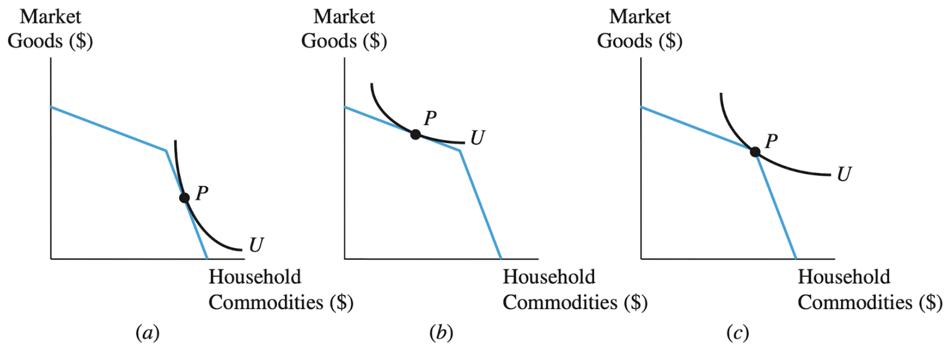

Figure 13: Various Household Allocations

Figure 14 captures what happens if Jack experiences an increase in wage or Jill experiences an increase in household productivity. In the first case, Jack shifts away from home production towards market work. In the second case, Jill shifts away from market work towards home production. In each case, these particular figures show complete specialization. In Panel (a), I show an extra bullet if there was no change in Jack's allocation between market work and home production. At this point, we can see that the household would only have to give up a little bit of household commodity to get much more in market goods, because of the newfound steepness of this portion of the household production frontier.

2.11 Policy Application: Welfare Programs and Work Incentives

This section considers the impact of various forms of welfare on work incentives. First, Figure 15 shows how a cash grant to those working zero hours could reduce a worker's supply of labor. In this figure, we see that the worker supplies G hours of labor at the considered wage, but if they receive $1,000$ for not working, they reach a higher indifference curve by taking the grant and not working at all. Thus, the worker with these example indifference curves facing this binary grant would choose not to work when they otherwise would've.

FIGURE 2-15 Increases in Wage Rate or Household Productivity Lead to Specialization
(*a*) An increase in Jack's wage moves the household from point *P* to *P′* and Jack specializes in the labor market.
(*b*) An increase in Jill's household productivity moves the household from point *P* to *P′* and Jill specializes in the household sector.

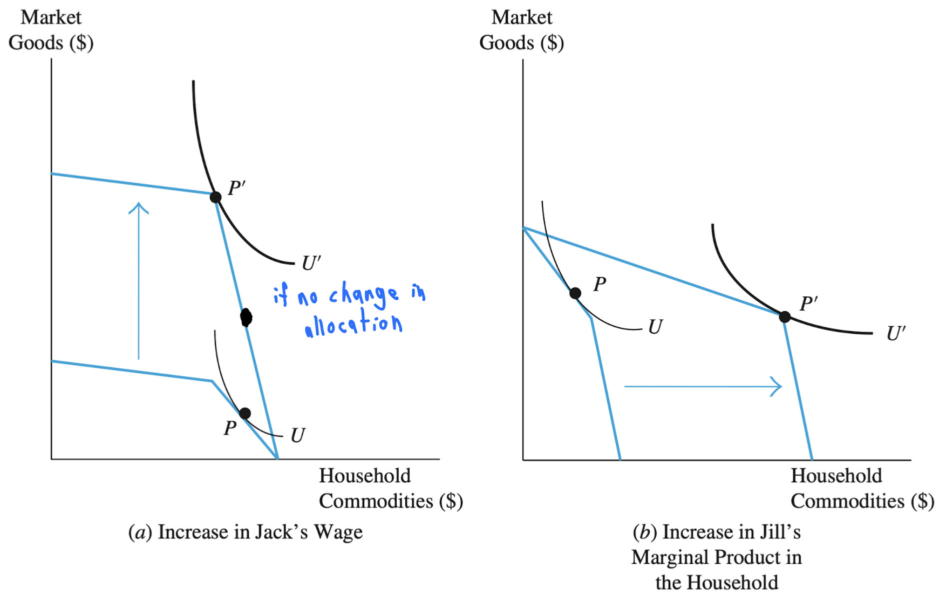

Figure 14: Impact of Wage or HH Production Changes on Household Production

FIGURE 2-16 Effect of a Cash Grant on Work Incentives
A take-it-or-leave-it cash grant of $1,000 per month moves the worker from point *P* to *G,* and she leaves the labor force.

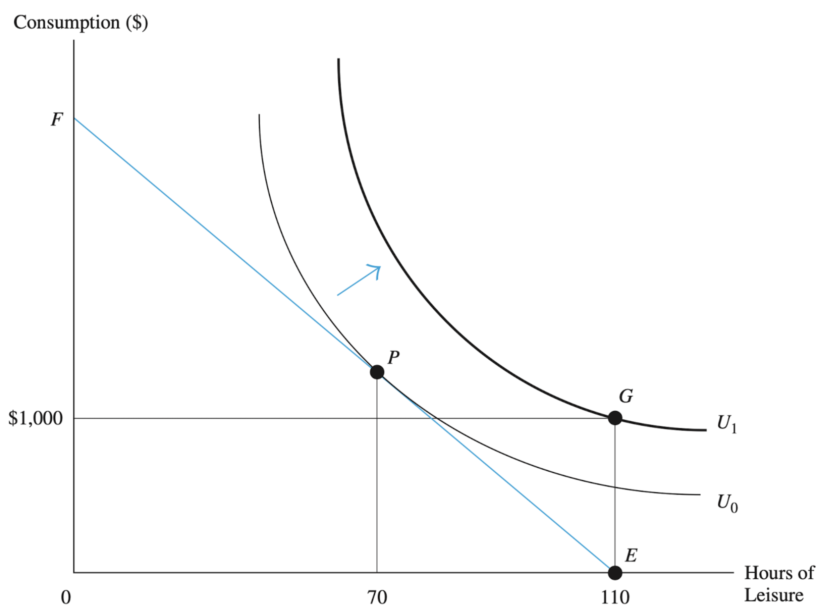

Figure 15: The Effect of a Cash Grant on Labor Supply

If we instead suppose that the worker receives a welfare benefit of $\$1,000$, which is then reduced by \$0.50 for every dollar earned in the market, we can think of this as a 50% tax on earnings for the first \$2,000 of earnings. An example of how this could play out for a worker is shown in Figure 16. As we can see, the policy shifts the opportunity set from being characterized by FE to HG, where FE has a slope of $-1w$ and HG has a slope of $-0.5w$. The agent's choice is characterized by P in the no-welfare scenario and R in the welfare scenario. In this example, the income effect moves the worker to Q, and the substitution effect then moves them to R. In this example, the income and substitution effects work in the same direction, since the income effect makes the worker work less when labor is a normal good, and the substitution effect makes the worker work less, given that leisure is less expensive because of the effective tax on earnings.

FIGURE 2-17 Effect of a Welfare Program on Hours of Work
The welfare program in budget line HG gives the worker a cash grant of \$1,000 and imposes a 50 percent tax on labor earnings. In the absence of welfare, the worker is at point *P*. The income effect resulting from the program moves the worker to point *Q*; the substitution effect moves the worker to point *R*. Both income and substitution effects reduce hours of work.

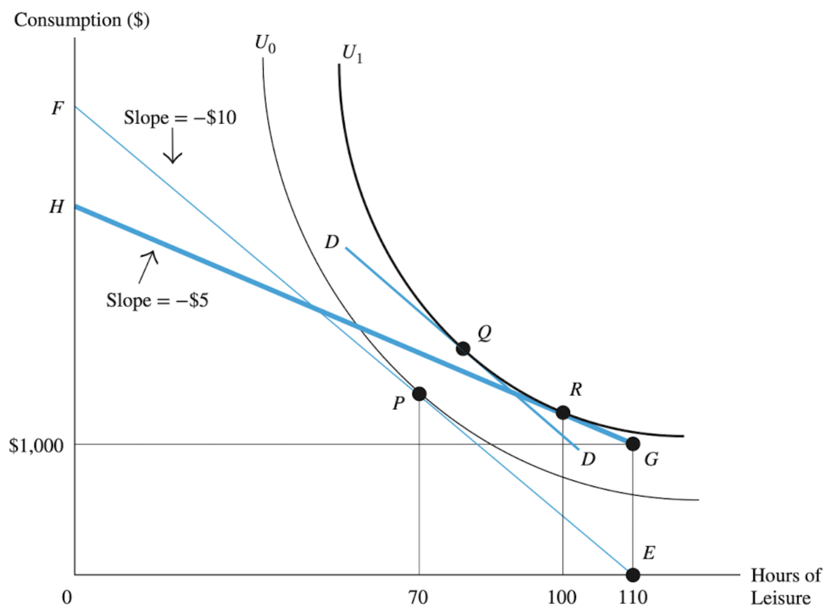

Figure 16: The Effect of a Welfare Program on Labor Supply

2.12 Policy Application: The Earned Income Tax Credit

The next example that we'll consider is the Earned Income Tax Credit (EITC). The EITC is a subsidy for low-income working families. The way that the EITC works is such that the workers earnings can be divided into four segments: (1) on earnings up to E_1, a subsidy rate of τ applies, so that the worker earns $(1+\tau)w$ per hour worked; (2) on earnings between E_1 and E_2, the worker earns w per hour worked; (3) on earnings between E_2 and E_3, the worker earns $(1-\rho)w$ per hour worked, where ρ is the phase-out rate;

and (4) on earnings above E_3, the worker earns w per hour worked. This is demonstrated graphically in Figure 17.

FIGURE 2-18 **The EITC and the Budget Line (Not Drawn to Scale)**
In the absence of the tax credit, the budget line is given by *FE*. The EITC grants the worker a credit of 40 percent on labor earnings as long she earns less than \$14,040. The credit is capped at \$5,616. The worker receives this amount as long as she earns between \$14,040 and \$18,340. The tax credit is then phased out gradually. The worker's net wage is 21.06 cents below her actual wage whenever she earns between \$18,340 and \$45,007.

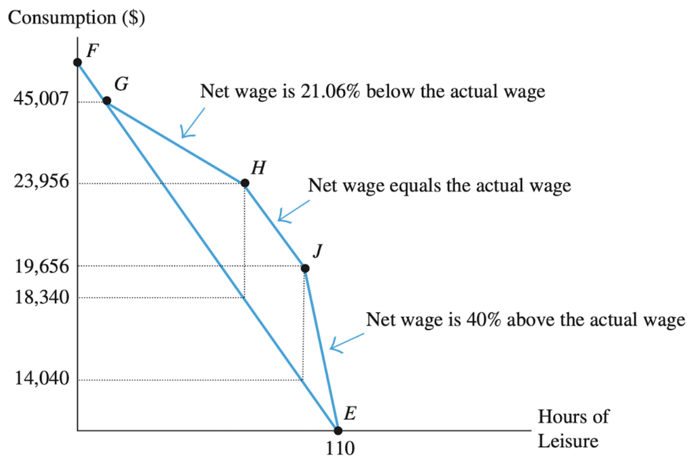

Figure 17: The Structure of the Earned Income Tax Credit

Figure 18 demonstrates some possible effects of the EITC on a worker's labor supply. Panel (a) demonstrates a worker who is brought into the labor force by the EITC. Panel (b) demonstrates a worker who reduces their hours worked due to the income effect of the EITC. Panel (c) demonstrates a worker who decreases their hours worked due to the income and substitution effects of the EITC.

FIGURE 2-19 **The Impact of the EITC on Labor Supply**

The EITC shifts the budget line, and will draw new workers into the labor market. In (*a*), the person enters the workforce by moving from point *P* to *R*. The impact of the EITC on the labor supply of persons who are already working is less clear. In the shifts illustrated in (*b*) and (*c*), the worker works fewer hours.

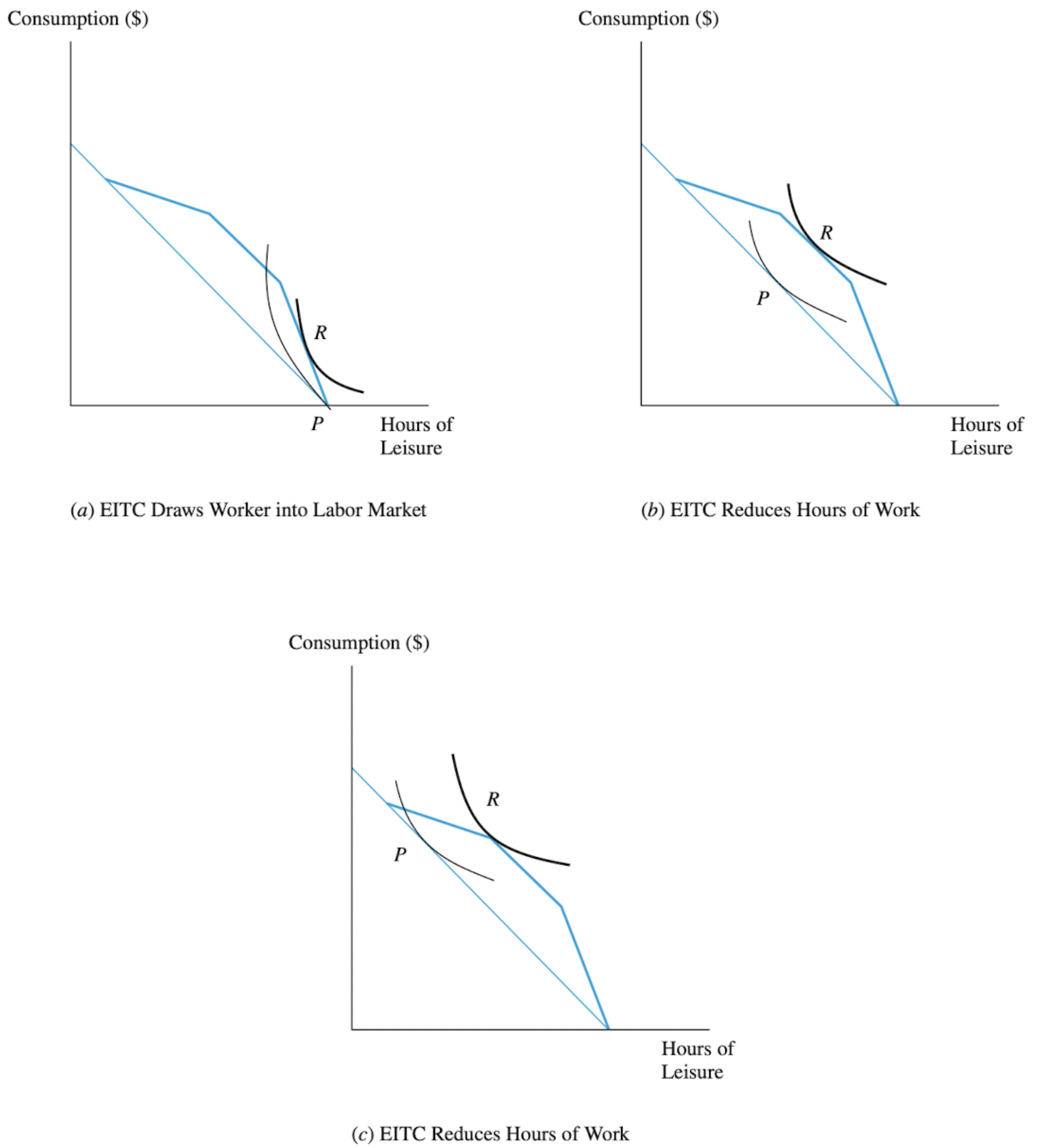

Figure 18: The Effects of the Earned Income Tax Credit

2.13 Labor Supply Over the Life Cycle

Consumption and leisure decisions are made over the course of the life cycle. Thus, workers can shift their labor to periods of relative higher productivity. This hypothesis is known as the "intertemporal substitution hypothesis." Figure 19 illustrates the typical earnings and hours worked patterns over the life cycle. Models of life cycle labor supply crucially depend on estimating the intertemporal labor supply elasticity. I won't include so much more discussion of this right now.

FIGURE 2-20 The Life Cycle Path of Wages and Hours for a Typical Worker
(*a*) The age-earnings profile of a typical worker rises rapidly when the worker is young, reaches a peak at around age 50, and then wages either stop growing or decline slightly. (*b*) The changing price of leisure implies that the worker will devote relatively more hours to the labor market when the wage is high and fewer hours when the wage is low.

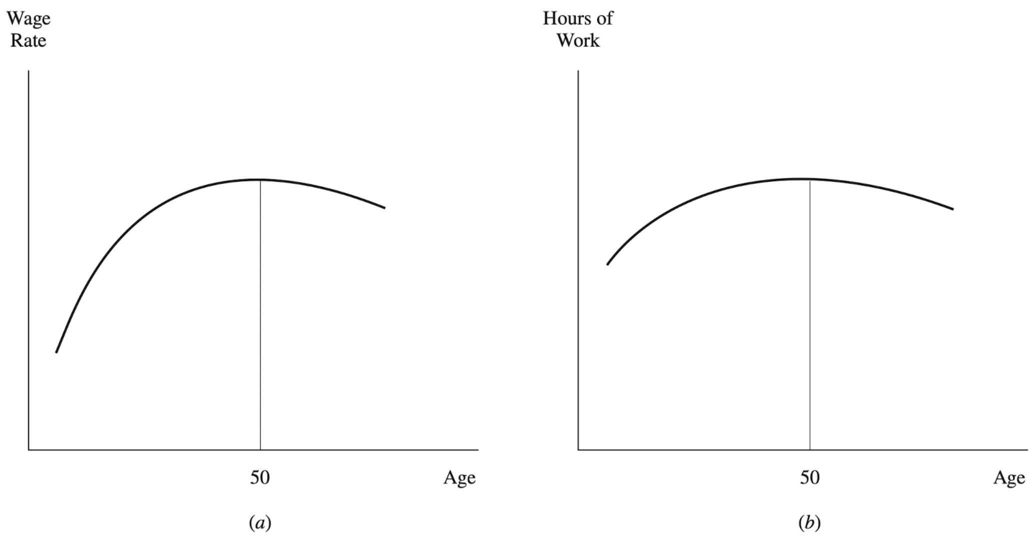

Figure 19: Wages and Hours Worked Over the Life Cycle

2.14 Policy Application: Disability Benefits and Labor Force Participation

The textbook then discusses some evidence that disability benefits reduce labor force participation. In particular, the author discusses some research suggesting that people who get denied disability benefits are more likely to work afterwards than people who are approved, even after attempting to sort through the endogeneity issues (e.g., by using an instrumental variable approach). I don't doubt that these results are true, but I'm not really sure what I'm meant to take away from it in a normative sense. Perhaps nothing.

3 Chapter 3: Labor Demand

3.1 Terms

- q: Quantity of output produced by the firm
- E: Employee-hours (labor input)
- K: Capital input (machines, land, etc.)
- MP_E: Marginal product of labor
- MP_K: Marginal product of capital
- p: Price of output
- w: Wage rate (price of labor)
- r: Rental rate (price of capital)
- VMP_E: Value of marginal product of labor
- VAP_E: Value of average product of labor
- C: Cost of production

3.2 The Firm's Production Function

The firm's production function describes the technology that the firm uses to produce its output. We will assume for the moment that the firm only requires two inputs for its production: labor in the form of employee-hours (E) and capital (K), which includes land, machines, and other physical inputs. We will write the firm's production function as:

$$q = f(E, K)$$

3.2.1 Marginal Product and Average Product

Definition D.10: Marginal Product of Labor

The marginal product of labor (MP_E) is the additional output produced as a result of hiring one more employee-hour, holding the amount of capital constant.

Definition D.11: Marginal Product of Capital

The marginal product of capital (MP_K) is the additional output produced as a result of using one more unit of capital, holding the amount of labor constant.

Definition D.12: Average Product of Labor

The average product of labor (AP_E) is the average output produced per employee-hour, holding the amount of capital constant.

$$AP_E = \frac{q}{E}$$

Definition D.13: Law of Diminishing Returns

This is the assumption that the marginal product of labor eventually declines.

Figure 20 shows an example of how we would calculate the marginal product, average product, value of marginal product, and value of average product for labor.

TABLE 3-1
Calculating the Marginal and Average Product of Labor (Holding Capital Constant)

Note: The calculations for the value of marginal product and the value of average product assume that the price of the output is $2.

Number of Workers Employed	Output (Units)	Marginal Product (Units)	Average Product (Units)	Value of Marginal Product ($)	Value of Average Product ($)
0	0	—	—	—	—
1	11	11	11.0	22	22.0
2	27	16	13.5	32	27.0
3	47	20	15.7	40	31.3
4	66	19	16.5	38	33.0
5	83	17	16.6	34	33.2
6	98	15	16.3	30	32.7
7	111	13	15.9	26	31.7
8	122	11	15.3	22	30.5
9	131	9	14.6	18	29.1
10	138	7	13.8	14	27.6

Figure 20: Marginal Product Table

Figure 21 shows the relationship between the marginal product and average product curves.

FIGURE 3-1 **The Total Product, Marginal Product, and Average Product Curves**
(*a*) The total product curve gives the relationship between output and the number of workers hired by the firm, holding capital fixed. (*b*) The marginal product curve gives the output produced by each additional worker and the average product curve gives the output per worker.

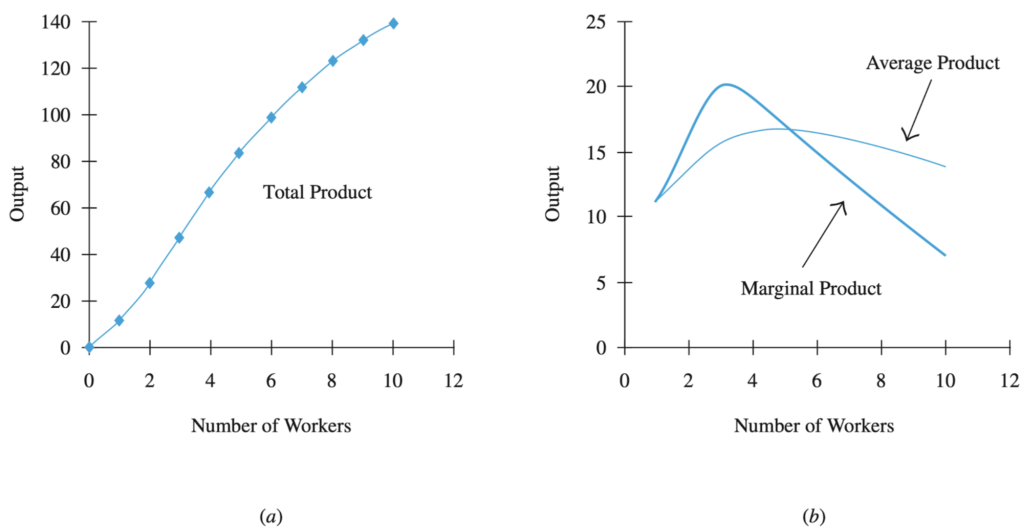

Figure 21: Marginal Product and Average Product Curves

3.2.2 Profit Maximization

The firm's profits are given by

$$\text{Profits} = pq - wE - rK$$

where

- p: Price of output
- w: Wage rate (price of labor)
- r: Rental rate (price of capital)

We assume that the firm is maximizing profits. We also, for the moment, take firms to be perfectly competitive, i.e., they cannot influence prices. Thus, given that these firms are price takers, they will maximize profits by choosing the optimal quantity of labor and capital.

Definition D.14: Perfectly Competitive Firm

We refer to a perfectly competitive firm as one that cannot influence prices.

3.3 The Short Run

Definition D.15: The Short Run

In our context, we define the short run to be a time horizon that is sufficiently short such that the firm cannot adjust its capital stock (K).

Definition D.16: Value of Marginal Product of Labor

The value of the marginal product of labor (VMP_E) is the additional revenue generated as a result of hiring one more employee-hour, holding the amount of capital constant.

$$VMP_E = p \cdot MP_E$$

Definition D.17: Value of Average Product of Labor

The value of the average product of labor (VAP_E) is the average revenue generated per employee-hour, holding the amount of capital constant.

$$VAP_E = p \cdot AP_E$$

3.3.1 How Many Workers Should the Firm Hire?

The firm will hire workers up to the point that the value of the marginal product of labor equals the wage rate and the value of the marginal product of labor is downward sloping. See Figure 22 for an example of VMP_E and VAP_E curves. In this figure, see the horizontal lines at 22 and 38. If the wage rate is 22, then the firm will hire 8 employees.[3] If the wage rate is 38, then the firm will hire 4 employees.

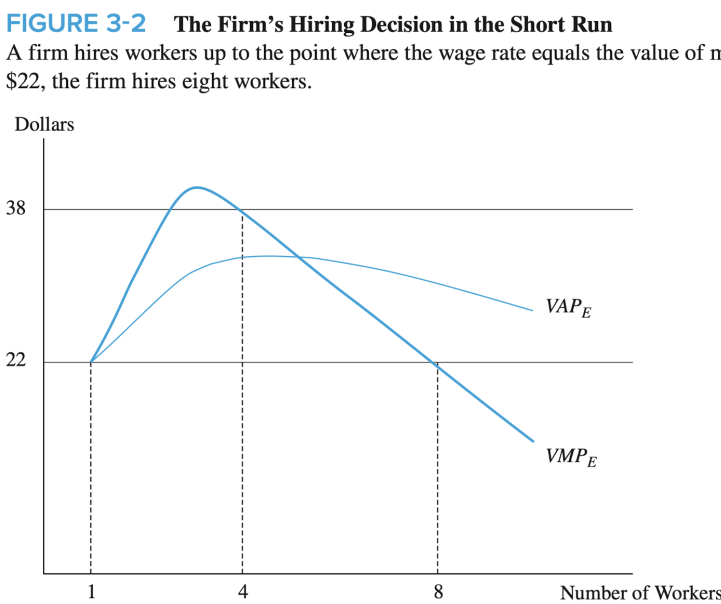

Figure 22: Short Run Hiring Decision

3.3.2 The Short-Run Labor Demand Curve for a Firm

With this setup, we can now derive the firm's short-run labor demand curve. In particular, the short-run labor-demand curve is given by the value of the marginal product of labor

[3] I think I'm kind of using employees and employee-hours as interchangeable concepts here.

curve. In particular, we can look at Figure 22, see at each value on the y-axis as a potential wage, draw a horizontal line to the point on the VMP_E curve that it intersects and is downward sloping, and then drop down to the x-axis to see how many employees the firm would hire at that wage. Doing this for all potential wages gives us the short-run labor demand curve, which is shown in Figure 23. Figure 23 also depicts what would happen if there was a shift in the VMP_E curve stemming from a rise in the price of the output. This would lead to more workers being hired, since it would raise the VMP_E curve to $VMP_{E'}$.

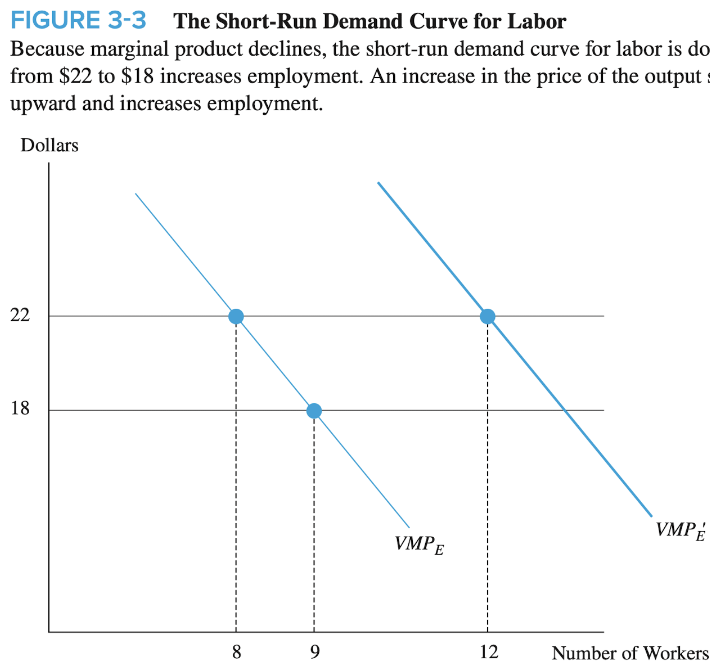

Figure 23: Short-Run Labor Demand Curve

Critical Point

It will be a recurring point to remember that the short-run labor demand curve for a firm is given by the VMP_E curve.

3.3.3 The Short-Run Labor Demand Curve in the Industry

The short-run labor demand curve in the industry is not simply the sum of the short-run labor demand curves of the individual firms, because, while one firm cannot influence prices, the industry as a whole can. If the industry increases production dramatically, the price of the output will fall, which lowers the VMP_E curve and hence demand. Thus, the short-run labor demand curve in the industry reflects this. See Figure 24 for an illustration.

FIGURE 3-4 The Short-Run Demand Curve for the Industry

Each firm in the industry hires 15 workers when the wage is $20. If the wage falls to $10, each firm hires 30 workers. If all firms expand, the output of the industry increases, reducing the price of the output and reducing the value of marginal product, so the labor demand curve of each individual firm shifts slightly to the left. At the lower price of $10, each firm hires only 28 workers. The industry demand curve is not given by the horizontal sum of the firms' demand curves (*DD*), but takes into account the impact of the industry's expansion on output price (*TT*).

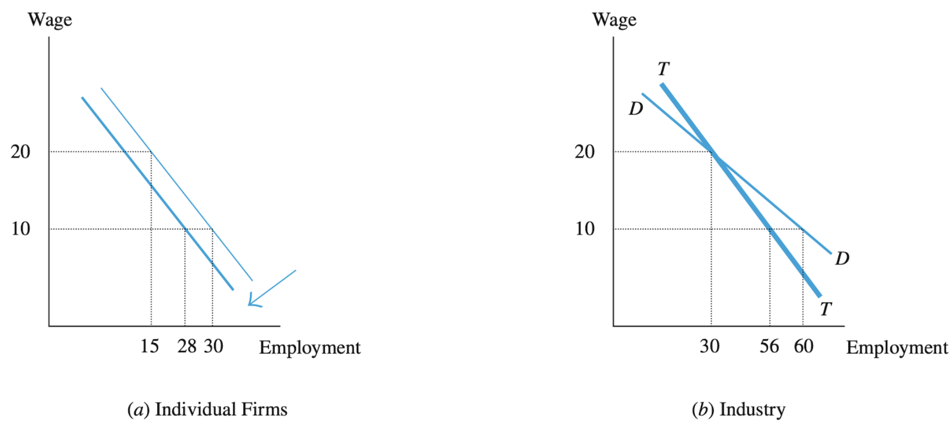

Figure 24: Short-Run Labor Demand Curve in the Industry

Definition D.18: Elasticity of Labor Demand

The elasticity of labor demand (δ_{SR}) is the percentage change in the quantity of labor demanded in the short run (E_{SR}) resulting from a percentage change in the wage rate.

$$\delta_{SR} = \frac{\Delta E_{SR}/E_{SR}}{\Delta w/w} = \frac{\Delta E_{SR}}{\Delta w} \cdot \frac{w}{E_{SR}}$$

3.3.4 An Alternative Interpretation of the Marginal Productivity Condition

An alternative way to express the point that the firm hires up to the point that the value of the marginal product of labor equals the wage rate is the following: The firm will hire up to the point that its marginal cost (MC)[4] equals its marginal revenue (MR). This interpretation is illustrated in Figure 25, where we show the marginal cost curve intersecting the marginal revenue curve (which in this case, is simply the price of the output, p). The optimal quantity of output is then given by q^*.

[4] That is, the marginal cost of producing another unit of the output.

FIGURE 3-5 **The Firm's Output Decision**
A profit-maximizing firm produces up to the point where the output price equals the marginal cost of production. This profit-maximizing condition is identical to the one stating that firms hire workers up to the point where the wage equals the value of marginal product.

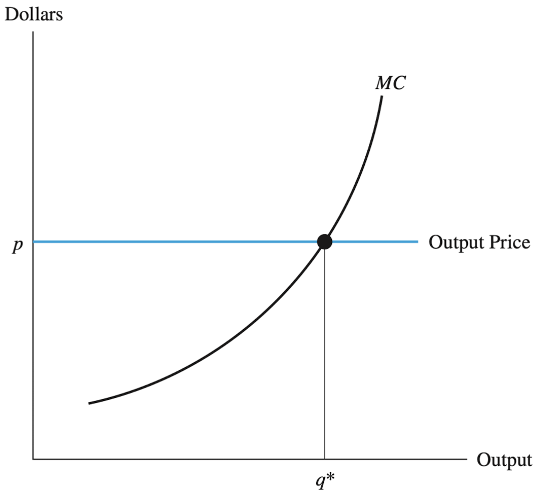

Figure 25: The Firm's Output Decision

We can note that the marginal cost of producing an extra unit of output is given by:

$$MC = w \times \frac{1}{MP_E}$$

That is, the wage scaled down by the marginal product of labor.[5]

Then, the profit-maximizing condition is that

$$w \times \frac{1}{MP_E} = p$$
$$\text{or } w = p \times MP_E = VMP_E$$

[5]You can think of this as saying something like: How much of a worker would you need to produce one more unit of output?

3.4 The Long Run

3.4.1 Isoquants

Definition D.19: Isoquant

An isoquant is a curve that captures the different combinations of labor and capital that produce the same level of output.

We demand the following properties of isoquants:

1. Isoquants must be downward sloping.
2. Isoquants do not intersect.
3. Higher isoquants are associated with higher levels of output.
4. Isoquants are convex to the origin.

See Figure 26 for an illustration of isoquants. Notice, as one would expect based on the definition, we are plotting the isoquants in K-E space.

FIGURE 3-6 **Isoquant Curves**
All capital–labor combinations along a single isoquant produce the same level of output. The input combinations at points X and Y produce q_0 units of output. Input combinations that lie on higher isoquants produce more output.

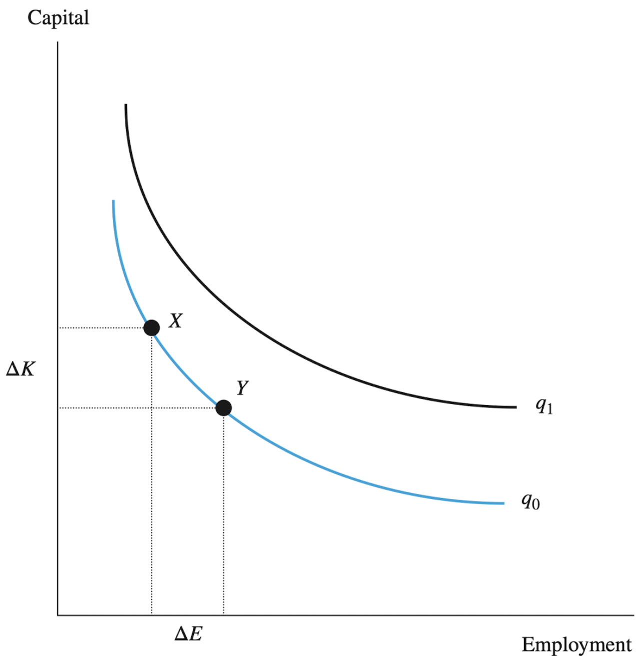

Figure 26: Isoquant Illustration

The slope of the isoquant is given by

$$\frac{\Delta K}{\Delta E} = -\frac{MP_E}{MP_K}$$

The LHS just captures the standard rise-over-run definition of slope. The RHS stems from the fact that, given the definition of an isoquant, it must be that:

$$(\Delta K)MP_K + (\Delta E)MP_E = 0$$

Thus, the slope of the isoquant is the negative ratio of the marginal products.

Definition D.20: Marginal Rate of Technical Substitution

The absolute value of the slope of the isoquant is referred to as the marginal rate of technical substitution (MRTS). That is, the MRTS is given by:

$$MRTS = \left|\frac{\Delta K}{\Delta E}\right| = \frac{MP_E}{MP_K}$$

Intuitively, this is telling us, if we increase labor by one unit, how much capital can we give up, while still keeping output constant.[a]

[a]The "one unit" framing isn't really right, since it's really about infinitesimal changes, but I'm being a little loose for the intuition.

3.4.2 Isocosts

The firm's cost of production is given by:

$$C = wE + rK$$

Definition D.21: Isocost Line

An isocost line is a curve that captures the different combinations of labor and capital that yield the same cost of production.

See Figure 27 for an illustration of isocost lines.

FIGURE 3-7 **Isocost Lines**
All capital–labor combinations along a single isocost curve are equally costly. Capital–labor combinations on a higher isocost curve are costlier. The slope of an isocost equals the ratio of input prices ($-w/r$).

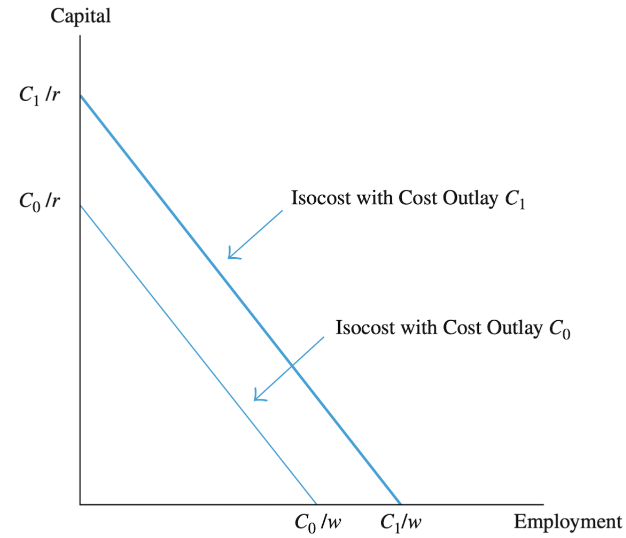

Figure 27: Isocost Lines

We can re-write the isocost line as:

$$K = \frac{C}{r} - \frac{w}{r}E$$

Thus, we can see that the slope of the isocost line is given by: $-\frac{w}{r}$. This is intuitive since it's telling us that the slope corresponds to telling us that an additional unit of labor costs $\frac{w}{r}$ units of capital.

3.4.3 Cost Minimization

Suppose that the firm is going to produce q_0 units of output. We can then ask: how can the firm produce q_0 at the lowest cost? The solution to this question is illustrated in Figure 28. The cost-minimizing solution is to produce at the point along the isocost line for q_0 which has a tangent isoquant.

FIGURE 3-8 **The Firm's Optimal Combination of Inputs**
A firm minimizes the cost of producing q_0 by using the capital–labor combination at point *P*, where the isoquant is tangent to the isocost. All other capital–labor combinations (such as those in points *A* and *B*) lie on a higher isocost.

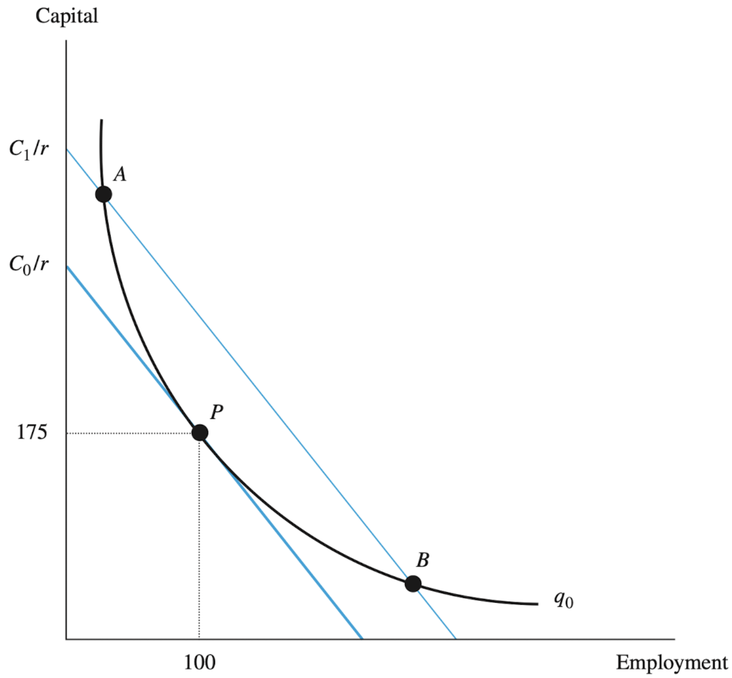

Figure 28: Cost Minimization

At the cost-minimizing solution, given the isoquant's tangency and the slope of each curve, we have:

$$\frac{MP_E}{MP_K} = \frac{w}{r}$$

That is the ratio of the marginal product of labor and capital equals the ratio of the cost of labor and capital.

The intuition may be more clearly seen by:

$$\frac{MP_E}{w} = \frac{MP_K}{r} \tag{1}$$

That is, the cost of the marginal product coming from labor and capital are equalized. It's logical that this is the optimal point, since if the cost of the marginal product from one was lower than the other, then the firm should re-allocate its spending toward that input.

3.4.3.1 Long-Run Profit Maximization

Note that the condition given by (1) is not the same as saying that firms maximize profit. This condition is only saying how firms minimize cost given a particular selection of output. Maximizing profit also requires choosing the optimal level of output.

Long-run profit maximization also necessitates that

$$w = p \times MP_E \quad \text{and} \quad r = p \times MP_K$$

which implies (1), but the reverse is not true.

Questions

Do we have a way of indicating the process of finding the profit-maximizing point on the isoquant-isocost diagram?

3.5 The Long-Run Demand Curve for Labor

Suppose that the wage for workers falls. How will firms respond in the long run? First, note that the worker's wage falling corresponds to a flattening of the isocost line (where the number of workers is on the x-axis).

Our first reaction to try to understand how the firms will re-allocate between capital and workers may be to draw Figure 29. However, this would be the wrong approach. In particular, this presupposes that the firm is going to hold costs fixed, but there is no reason to believe that this is the case.

FIGURE 3-9 The Impact of a Wage Reduction, Holding Constant Initial Cost Outlay at C_0
A wage reduction flattens the isocost curve. If the firm were to hold the initial cost outlay constant at C_0 dollars, the isocost would rotate around C_0 and the firm would move from point P to point R. A profit-maximizing firm, however, will not generally want to hold the cost outlay constant when the wage changes.

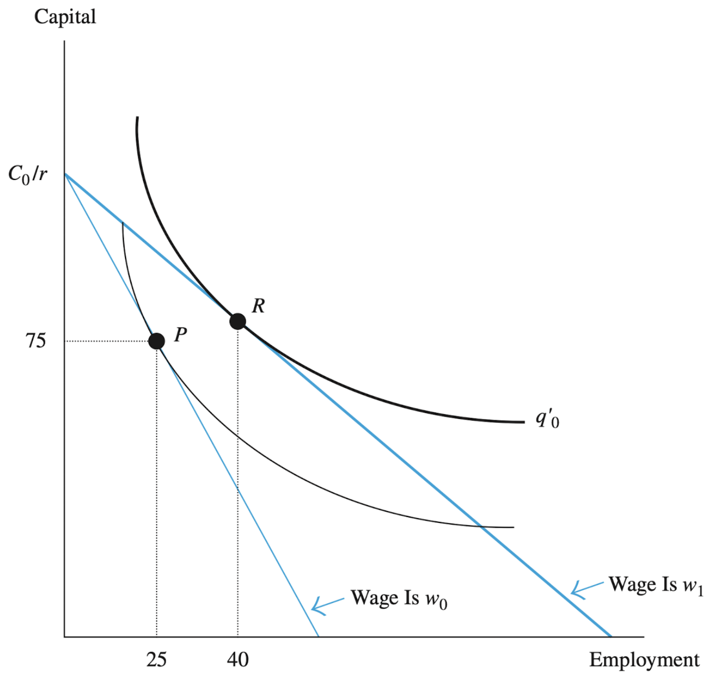

Figure 29: The Impact of a Wage Decrease if C_0 is Fixed

3.5.1 Will the Firm Expand if the Wage Falls?

Accommodating the reality that the cost may change, we can look to Figure 30 for an example of what might happen. Specifically, the fact that the marginal cost curve has shifted down encourages the firm to increase production. (See Panel (a).) Then, we can look at the relevant isoquant and isocost in Panel (b) to see what's happened to employment and capital usage. In this example, we display both as having increased. The important takeaway here, though, is primarily just that the firm is producing at a different level of cost.

FIGURE 3-10 **Impact of Wage Reduction on Output and Employment of a Profit-Maximizing Firm**
(*a*) A wage cut reduces the marginal cost of production and encourages the firm to expand (from producing 100 to 150 units). (*b*) The firm moves from point *P* to point *R*, increasing the number of workers hired from 25 to 50.

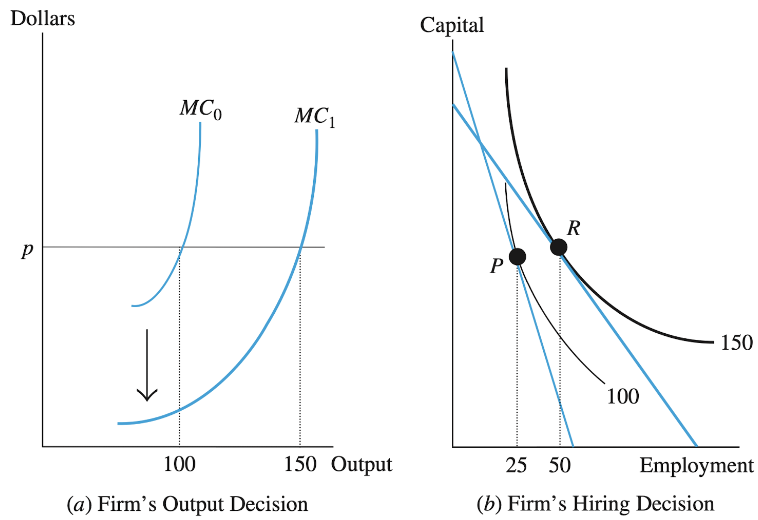

Figure 30: The Impact of a Wage Decrease if C_0 is Not Fixed

3.5.2 Substitution and Scale Effects

This section is analogous to our discussion of income and substitution effects for consumers. Here, we consider two effects that accompany a change in an input price for a firm: the scale effect and the substitution effect.

Definition D.22: Scale Effect

The scale effect refers to what happens to the firm's demands for its inputs as production increases, holding input prices constant.

Definition D.23: Substitution Effect

The substitution effect refers to what happens to the firm's demands for its inputs as input prices change, holding production constant.

In Figure 31, we illustrate these two effects in the context of a wage decrease. Prior to the wage change, the firm chooses employment and capital levels corresponding to P. If the production were shifted to the new level (150, in this example), holding input prices constant, the firm would choose the point Q, which corresponds to an increase in both capital and employment. This is the scale effect. The increase in both inputs holds as long as both are "normal inputs." However, if we now adjust the input prices to reflect the wage decrease, holding production constant at this new level, the firm would produce at point R. This reflects a decline in capital usage, relative to Q, and an increase in employment. This is the substitution effect. The directional effect of the substitution effect is not ambiguous.

FIGURE 3-11 Substitution and Scale Effects

A wage cut generates substitution and scale effects. The scale effect (the move from point *P* to point *Q*) encourages the firm to expand, increasing the firm's employment. The substitution effect (from *Q* to *R*) encourages the firm to use a more labor-intensive method of production, further increasing employment.

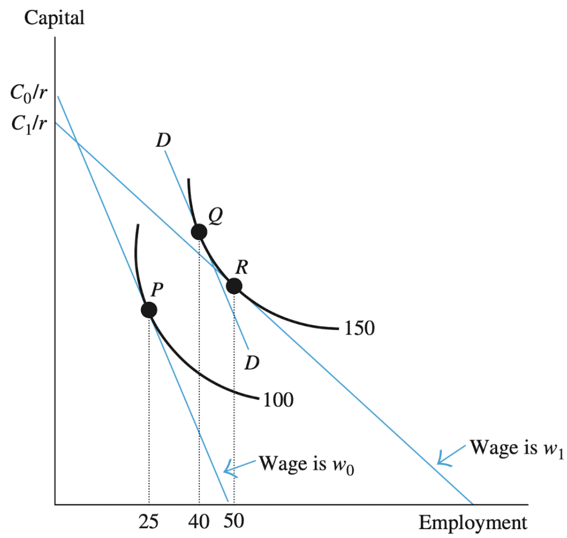

Figure 31: Substitution and Scale Effects

Definition D.24: Long-Run Elasticity of Labor Demand

The long-run elasticity of labor demand (δ_{LR}) is the percentage change in the quantity of labor demanded in the long run (E_{LR}) resulting from a percentage change in the wage rate.

$$\delta_{LR} = \frac{\Delta E_{LR}/E_{LR}}{\Delta w/w} = \frac{\Delta E_{LR}}{\Delta w} \cdot \frac{w}{E_{LR}}$$

In general, long-run labor demand will be more elastic than short-run labor demand. This dynamic is because in the long run, firms can adjust both capital and labor, so they can substitute between the two inputs.

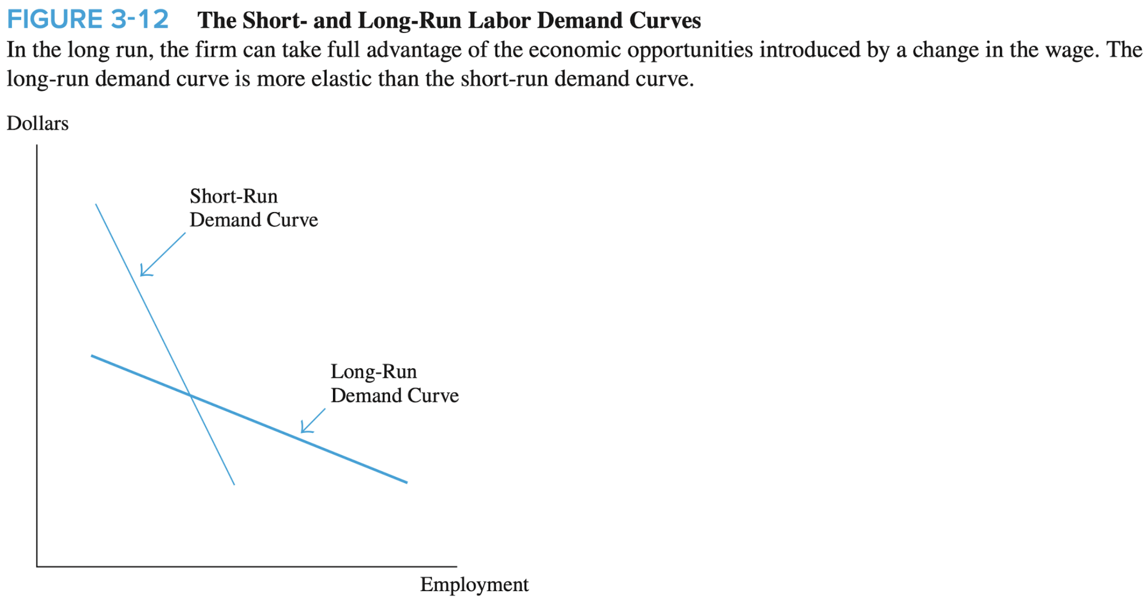

Figure 32: Long-Run vs Short-Run Elasticity of Labor Demand

3.5.3 Estimates of the Elasticity of Labor Demand

The author suggests that the short-run elasticity of labor demand is in the range of -0.4 and -0.5. That is, a 10% increase in wages leads to a 4-5% decrease in employment. He suggests that the long-run elasticity is estimated to be around -1%, i.e., a 10% increase in wages leads to a 10% decrease in employment.

3.6 The Elasticity of Substitution

"The size of the firm's substitution effect depends on the curvature of the isoquant."

Definition D.25: Perfect Substitute

If the isoquants are linear, then the inputs are perfect substitutes.

In the case of perfect substitutes, firms will simply use the cheapest input.

Definition D.26: Perfect Complement

If the isoquants are right angles, then the inputs are perfect complements.

In the case of perfect complements, firms will use inputs in fixed proportions.

See Figure 33 for an illustration of perfect substitutes and perfect complements.

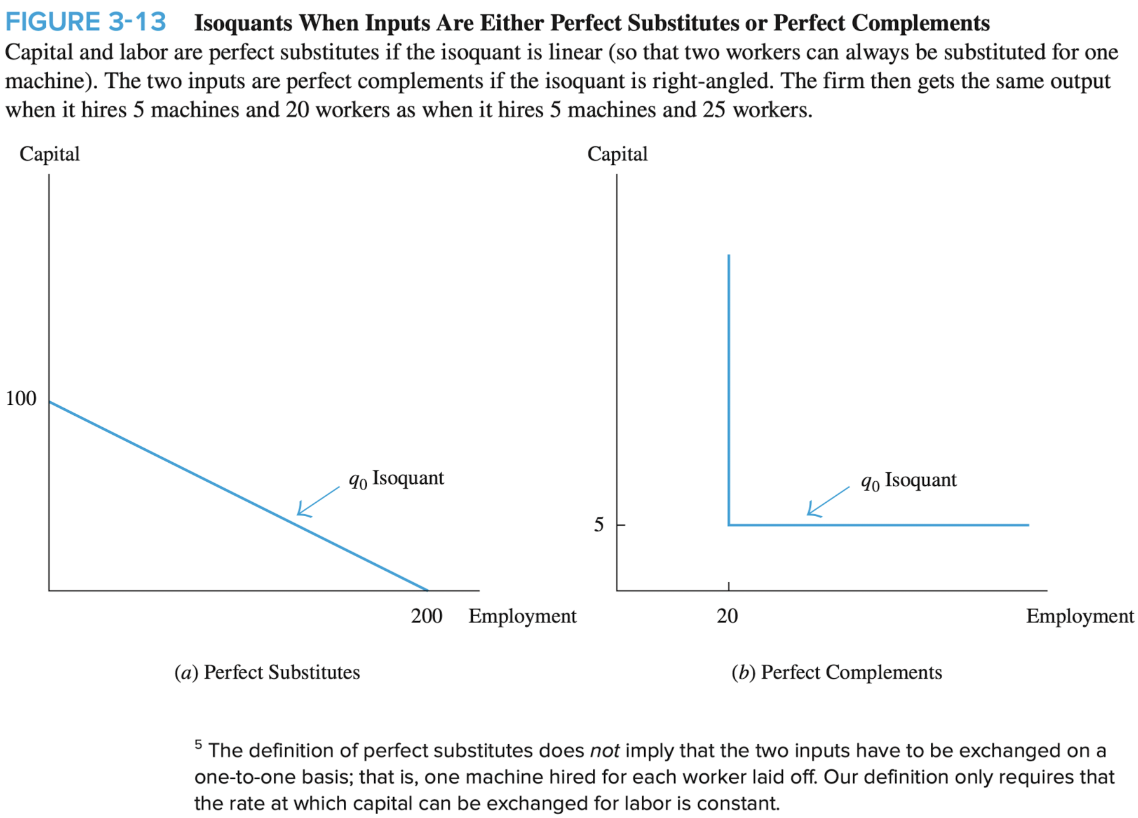

Figure 33: Perfect Substitutes and Perfect Complements

Definition D.27: Elasticity of Subsitution

The elasticity of substitution (σ) is the percentage change in the capital-labor ratio (K/L) resulting from a percentage change in the capital-labor price ratio (w/r).

$$\sigma = \frac{\text{Percent change in } K/L}{\text{Percent change in } w/r}$$

3.7 What Makes Labor Demand Elastic?

Marshall's rules of derived demand describe situations in which we would expect labor demand to be more elastic:

1. "Labor demand is more elastic the greater the elasticity of substitution."

2. "Labor demand is more elastic the greater the elasticity of demand for the output."
3. "Labor demand is more elastic the greater labor's share in total costs."
4. "The demand for labor is more elastic the greater the supply elasticity of other factors of production."

3.8 Factor Demand with Many Inputs

We now consider extending the theory we've developed so far to more inputs than simply labor and capital.

Suppose the firm's production function is instead given by:

$$q = f(x_1, x_2, x_3, \ldots, x_n)$$

where each x_i is a different input. We denote the marginal product of input i as MP_i, and the price of input i as w_i.

Under this setup, we still have the result that the ith input is purchased up to the point that:

$$w_i = p \times MP_i$$

Definition D.28: Cross-Elasticity of Factor Demand

As a measure of how the demand for one input responds to a change in the price of another input, we consider the cross-elasticity of factor demand:

$$\delta_{ij} = \frac{\%\Delta x_i}{\%\Delta w_j}$$

If the cross-elasticity is positive, then we say that the two inputs are substitutes. If the cross-elasticity is negative, then we say that the two inputs are complements.

Definition D.29: Capital-Skill Complementarity Hypothesis

The capital-skill complementarity hypothesis is the hypothesis that unskilled labor and capital are substitutes, while skilled labor and capital are complements.

3.9 Overview of Labor Market Equilibrium

We will cover it more fully in the next chapter, but now we will provide a brief overview of the idea of labor market equilibrium. The main idea is that the labor market occurs at the intersection of the labor supply and labor demand curves.

Figure 34 shows an example of labor market equilibrium at (E^*, w^*). At this point, in contrast to any other wage, the number of workers who want to work equals the number of workers that firms want to hire.

FIGURE 3-15 **Wage and Employment Determination in a Competitive Market**
In a competitive labor market, equilibrium is attained where supply equals demand. The equilibrium wage is w^* and E^* workers are employed.

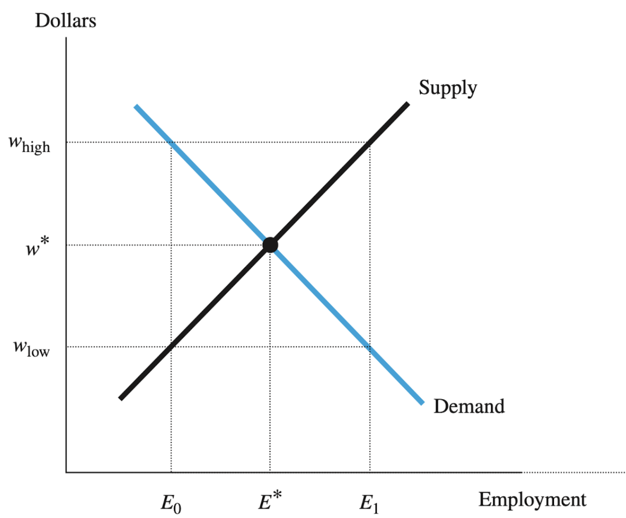

Figure 34: Labor Market Equilibrium Example

3.10 Rosie the Riveter as an Instrumental Variable

The author notes that much of modern labor economics research is focused on estimating labor supply and demand curves for various groups.

One method for doing so is to identify an instrumental variable (IV) that shifts only either supply or demand. This shift can then be used to trace out the other. The logic of this approach is visualized in Figure 35. Suppose the supply and demand curves start at S_0 and D_0, respectively. Then, consider three scenarios:

- S_0 moves to S_1, while D_0 stays fixed: Then the equilibrium moves from P to Q, and we can trace out the demand curve between these two points.

- D_0 moves to D_1, while S_0 stays fixed: Then the equilibrium moves from P to the unlabeled intersection of S_0 and D_1. Then, we can trace out the supply curve between these two points.
- S_0 moves to S_1 and D_0 moves to D_1: Then the equilibrium moves from P to R, and we can't say anything about either curve.

FIGURE 3-16 **Shifts in Supply and Demand Curves Generate Observed Wage and Employment Data**
The market is initially in equilibrium at P, and we observe wage w_0 and employment E_0. If only the supply curve shifts, we can observe w_1 and E_1, and the available data lets us to trace out the labor demand curve. But if both the supply and demand curves shift, we observe w_2 and E_2, and the available data trace out the curve ZZ, which does not provide any information about either the supply or the demand curve.

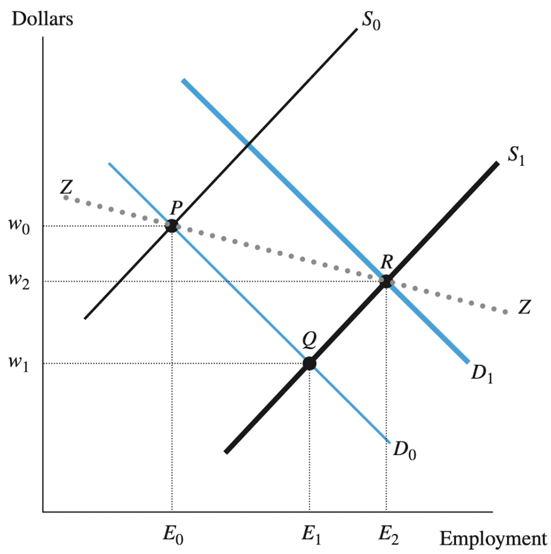

Figure 35: Instrumental Variable Approach

The author then describes one paper that uses this approach in the context of WWII, in which geographic differences in the mobilization rate of men was used as an instrument for female labor supply.

3.11 Policy Application: The Minimum Wage

We start by studying the effect of the minimum wage in a perfectly competitive labor market; we will discuss its effects under monospony in a future chapter. Figure 36 will be our reference point for this discussion. Prior to the minimum wage, the labor market is in equilibrium at (E^*, w^*). Then, suppose that a minimum wage is imposed at $\bar{w}$. Now, the labor supply is E_S, while the labor demand is $\bar{E}$. Thus, there are $E^* - \bar{E}$ fewer workers employed, and $E_S - E^*$ workers who want to work but cannot find jobs (i.e., who are "unemployed").

FIGURE 3-19 **The Impact of the Minimum Wage on Employment**
A minimum wage set at $\overline{w}$ forces employees to cut employment (from E^* to $-\overline{E}$). The higher wage also encourages ($E_S - E^*$) additional workers to enter the market. The minimum wage, therefore, creates unemployment.

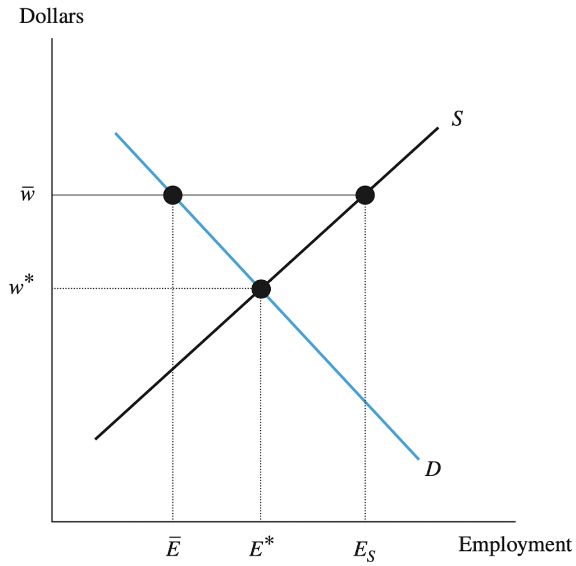

Figure 36: Minimum Wage Effects

The author then discusses compliance with minimum wage and how it relates to covered and uncovered sectors, but I won't discuss these here at the moment. The author then gives a lengthy discussion of the evidence and debates surrounding the effect of the minimum wage on employment. Again, I will not relay it all here, but it's available in the book for review.