Chapter 2

Labor Supply

It's true hard work never killed anybody, but I figure, why take the chance?

—*Ronald Reagan*

Each of us must decide whether to work and, once employed, how many hours to work. At any point in time, the economywide labor supply is given by adding the work choices made by each person in the population. Total labor supply also depends on the fertility decisions made by earlier generations (which determine the size of the current population).

The economic and social consequences of these decisions vary dramatically over time. In 1948, 84 percent of American men and 31 percent of American women aged 16 or over worked. By 2010, the proportion of working men had declined to 64 percent, whereas the proportion of working women had risen to 54 percent. Over the same period, the length of the average workweek in a private-sector production job fell from 40 to 34 hours.[1] These labor supply trends have surely altered the nature of the American family as well as greatly affected the economy's productive capacity.

This chapter develops the framework that economists use to study labor supply decisions. In this framework, individuals seek to maximize their well-being by consuming goods (such as fancy cars and nice homes) and leisure. Goods have to be purchased in the marketplace. Because most of us are not independently wealthy, we must work in order to earn the cash required to buy the desired goods. The economic trade-off is clear: If we do not work, we can consume a lot of leisure, but we have to do without the goods and services that make life more enjoyable. If we do work, we will be able to afford many of these goods and services, but we must give up some of our valuable leisure time.

The model of labor-leisure choice isolates the person's wage rate and income as the key economic variables that guide the allocation of time between the labor market and leisure activities. In this chapter, we first use the framework to analyze "static" labor supply

---

[1] These statistics were obtained from the U.S. Bureau of Labor Statistics Web site: www.bls.gov/data/home.htm.

decisions, the decisions that affect a person's labor supply at a point in time. We will also extend the basic model to explore how the timing of leisure activities changes over the life cycle.

This economic framework not only helps us understand why women's work propensities rose and hours of work declined, but also allows us to address a number of questions with important policy and social consequences. For example, do welfare programs reduce incentives to work? Does a cut in the income tax rate increase hours of work? And what factors explain the rapid growth in the number of women who choose to participate in the labor market?

2-1   Measuring the Labor Force

On the first Friday of every month, the Bureau of Labor Statistics (BLS) releases its estimate of the unemployment rate for the previous month. The unemployment rate statistic is widely regarded as a measure of the overall health of the U.S. economy. In fact, the media often interpret the minor month-to-month blips in the unemployment rate as a sign of either a precipitous decline in economic activity or a surging recovery.

The unemployment rate is tabulated from the responses to a monthly BLS survey called the *Current Population Survey* (CPS). In this survey, nearly 50,000 households are questioned about their work activities during a particular week of the month (that week is called the reference week). Almost everything we know about the trends in the U.S. labor force comes from tabulations of CPS data. The survey instrument used by the CPS also has influenced the development of surveys in many other countries. In view of the importance of this survey in the calculation of labor force statistics both in the United States and abroad, it is useful to review the various definitions of labor force activities that are routinely used by the BLS to generate its statistics.

The CPS classifies all persons aged 16 or older into one of three categories: the *employed,* the *unemployed,* and the residual group that is said to be *out of the labor force.* To be employed, a worker must have been at a job with pay for at least 1 hour or worked at least 15 hours on a nonpaid job (such as the family farm). To be unemployed, a worker must either be on a temporary layoff from a job or have no job but be actively looking for work in the four-week period prior to the reference week.

Let E be the number of persons employed and U the number of persons unemployed. A person participates in the **labor force** if he or she is either employed or unemployed. The size of the labor force (LF) is given by

$$LF = E + U \qquad\qquad \text{(2-1)}$$

Note that the vast majority of employed persons (those who work at a job with pay) are counted as being in the labor force regardless of how many hours they work. The size of the labor force, therefore, does not say anything about the "intensity" of work.

The **labor force participation rate** gives the fraction of the population (P) that is in the labor force and is defined by

$$\text{Labor force participation rate} = \frac{LF}{P} \qquad\qquad \text{(2-2)}$$

The **employment rate** gives the fraction of the population that is employed, or

$$\text{Employment rate} = \frac{E}{P} \qquad \textbf{(2-3)}$$

Finally, the **unemployment rate** gives the fraction of labor force participants who are unemployed:

$$\text{Unemployment rate} = \frac{U}{LF} \qquad \textbf{(2-4)}$$

The Hidden Unemployed

The BLS calculates an unemployment rate based on a subjective measure of what it means to be unemployed. To be considered unemployed, a person must either be on temporary layoff or claim that he or she has "actively looked for work" in the past four weeks. Persons who have given up and stopped looking for work are not counted as unemployed, but are considered to be "out of the labor force." At the same time, some persons who have little intention of working at the present time may claim to be "actively looking" for a job in order to qualify for unemployment benefits.

The unemployment statistics, therefore, can be interpreted in different ways. During the severe recession that began in 2009, for instance, it is often argued that the official unemployment rate (that is, the BLS statistic) understates the depths of the recession and economic hardships. Because it is so hard to find work, many laid-off workers have become discouraged with their futile job search activity, dropped out of the labor market, and stopped being *counted as* unemployed. It is then argued that this army of **hidden unemployed** should be added to the pool of unemployed workers so that the unemployment problem is significantly worse than it appeared from the BLS data.[2]

Some analysts have argued that a more objective measure of aggregate economic activity may be given by the employment rate. The employment rate simply indicates the fraction of the population at a job. This statistic has the obvious drawback that it lumps together persons who say they are unemployed with persons who are classified as being out of the labor force. Although the latter group includes some of the hidden unemployed, it also includes many individuals who have little intention of working at the present time (for example, retirees, women with small children, and students enrolled in school).

A decrease in the employment rate could then be attributed to either increases in unemployment or unrelated increases in fertility or school enrollment rates. It is far from clear, therefore, that the employment rate provides a better measure of fluctuations in economic activity than the unemployment rate. We shall return to some of the questions raised by the ambiguity in the interpretation of the BLS labor force statistics in Chapter 12.

---

[2] If one included the hidden unemployed as measured by the BLS (which counts persons who are out of the labor force because they are "discouraged over job prospects") as well as persons who are only "marginally attached" to the labor force, the unemployment rate in March 2011 would have increased from the official 8.8 percent to 15.7 percent.

2-2 Basic Facts about Labor Supply

This section summarizes some of the key trends in labor supply in the United States.[3] These facts have motivated much of the research on labor supply conducted in the past three decades. Table 2-1 documents the historical trends in the labor force participation rate of men. There was a slight fall in the labor force participation rates of men in the twentieth century, from 80 percent in 1900 to 72 percent by 2009. The decline is particularly steep for men near or above age 65, as more men choose to retire earlier. The labor force participation rate of men aged 45 to 64, for example, declined by 11 percentage points between 1950 and 2009, while the participation rate of men over 65 declined from 46 to 22 percent over the same period. Moreover, the labor force participation rate of men in their prime working years (ages 25 to 44) also declined, from 97 percent in 1950 to 91 percent in 2009. Note, however, that the labor force participation rate of men in their retirement years has begun to increase in the past 20 years.[4]

As Table 2-2 shows, there also has been a huge increase in the labor force participation rate of women. At the beginning of the century, only 21 percent of women were in the labor force. As late as 1950, even after the social and economic disruptions caused by two world wars and the Great Depression, only 29 percent of women were in the labor force. During the past 50 years, however, the labor force participation rate of women has increased dramatically. By 2009, almost 60 percent of all women were in the labor force.

TABLE 2-1 **Labor Force Participation Rates of Men, 1900–2009**

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	74.7	93.1	78.3	17.5
2009	72.0	91.0	80.8	21.9

[3] For more detailed discussions of the trends in labor supply in the United States and in other countries, see John H. Pencavel, "Labor Supply of Men: A Survey," in Orley C. Ashenfelter and Richard Layard, editors, *Handbook of Labor Economics,* vol. 1, Amsterdam: Elsevier, 1986, pp. 3–102; and Mark R. Killingsworth and James J. Heckman, "Female Labor Supply: A Survey," in ibid., pp. 103–204. See also Mark R. Killingsworth, *Labor Supply*, Cambridge: Cambridge University Press, 1983.

[4] See Tammy Schirle, "Why Have the Labor Force Participation Rates of Older Men Increased since the Mid-1990s?" *Journal of Labor Economics* 26 (October 2008): 549–594.

TABLE 2-2 **Labor Force Participation Rates of Women, 1900–2009**

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	60.2	69.0	61.3	49.4
2009	59.2	64.2	61.4	49.3

It is worth noting that the increase in female labor force participation was particularly steep among married women. Their labor force participation rate almost doubled in recent decades, from 32 percent in 1960 to 61.4 percent in 2009.

These dramatic shifts in labor force participation rates were accompanied by a sizable decline in average hours of work per week. Figure 2-1 shows that the typical person employed in production worked 55 hours per week in 1900, 40 hours in 1940, and just under 34 hours in 2010.[5]

There exist sizable differences in the various dimensions of labor supply across demographic groups at a particular point in time. As Table 2-3 shows, men not only have larger participation rates than women, but are also less likely to be employed in part-time jobs. Only 6 percent of working men are in part-time jobs, as compared to 16 percent of working women. The table also documents a strong positive correlation between labor supply and educational attainment for both men and women. In 2010, 92 percent of male college graduates and 80 percent of female college graduates were in the labor force, as compared to only 74 and 48 percent of male and female high school dropouts, respectively. There are also racial differences in labor supply, with white men having higher participation rates and working more hours than black men.

Finally, the decline in average weekly hours of work shown in Figure 2-1 was accompanied by a substantial increase in the number of hours that both men and women devote to leisure activities. It has been estimated that the number of weekly leisure hours increased by 6.2 hours for men and 4.9 hours for women between 1965 and 2003.[6]

---

[5] An interesting study of the trends in the length of the workday is given by Dora L. Costa, "The Wage and the Length of the Work Day: From the 1890s to 1991," *Journal of Labor Economics* 18 (January 2000): 156–181. She finds that low-wage workers had the longest workday at the beginning of the twentieth century. By the 1990s, however, this trend was reversed and high-wage workers had the longest workday. See also Peter Kuhn and Fernando Lozano, "The Expanding Workweek? Understanding Trends in Long Work Hours among U.S. Men, 1979–2006," *Journal of Labor Economics* 26 (April 2008): 311–343.

[6] Mark Agular and Erik Hurst, "Measuring Trends in Leisure: Allocation of Time over Five Decades," *Quarterly Journal of Economics* 122 (August 2007): 969–1006.

FIGURE 2-1   Average Weekly Hours of Work of Production Workers, 1900–2010

Sources: The pre-1947 data are drawn from Ethel Jones, "New Estimates of Hours of Work per Week and Hourly Earnings, 1900–1957," *Review of Economics and Statistics* 45 (November 1963): 374–385. Beginning in 1947, the data are drawn from U.S. Department of Labor, Bureau of Labor Statistics, *Employment, Hours, and Earnings from the Current Employment Statistics Survey,* "Table B-7. Average Weekly Hours of Production or Nonsupervisory Workers on Private Nonfarm Payrolls by Industry Sector and Selected Industry Detail": www.bls.gov/ces/cesbtabs.htm.
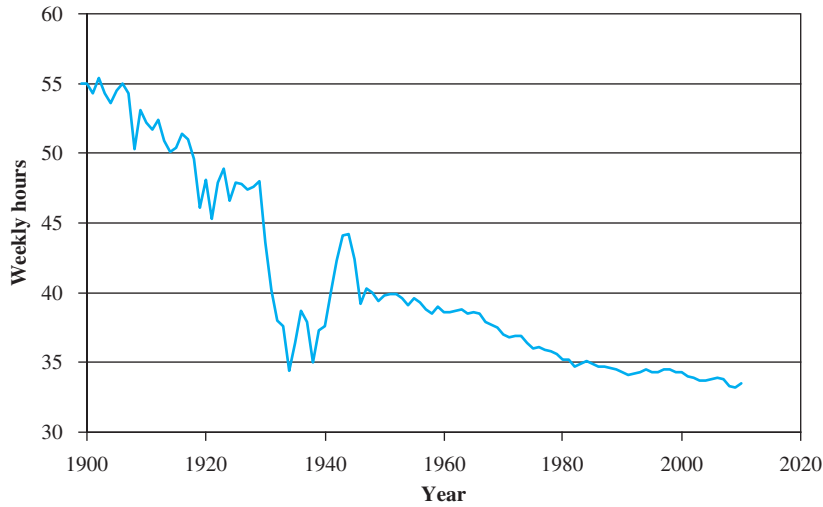
TABLE 2-3   Labor Supply in the United States, 2010 (persons aged 25–64)

Source: U.S. Bureau of Labor Statistics, *Current Population Survey*, March 2010. The average number of hours worked is calculated in the subsample of workers. The percent of workers in part-time jobs refers to the proportion working fewer than 30 hours per week.

	Labor Force Participation Rate		Annual Hours of Work		Percent of Workers in Part-Time Jobs	
	Men	Women	Men	Women	Men	Women
All persons	85.4	72.4	2,031	1,797	5.8	15.5
Educational attainment:						
Less than 12 years	74.0	48.2	1,763	1,617	9.4	18.5
12 years	83.1	68.2	1,949	1,755	5.8	15.8
13–15 years	85.6	75.0	2,030	1,771	6.2	16.3
16 years or more	91.6	80.4	2,182	1,878	4.6	14.1
Age:						
25–34	89.9	74.5	1,930	1,749	7.0	14.4
35–44	91.6	76.1	2,084	1,798	4.3	15.8
45–54	86.9	76.5	2,089	1,853	4.6	14.2
55–64	70.5	60.8	2,015	1,777	8.0	18.6
Race:						
White	86.2	74.0	2,079	1,799	5.3	16.6
Black	77.2	71.9	1,934	1,832	6.3	10.9
Hispanic	87.4	65.9	1,879	1,739	7.5	14.9

The data presented in this section provide the basic "stylized facts" that have motivated much of the work on the economics of labor supply. As we will see below, the evidence suggests that changes in the economic environment—particularly in wage rates and incomes—can account for many of the observed shifts in labor supply.

2-3   The Worker's Preferences

The framework that economists typically use to analyze labor supply behavior is called the **neoclassical model of labor-leisure choice.** This model isolates the factors that determine whether a particular person works and, if so, how many hours she chooses to work. By isolating these key factors, we can tell a simple "story" that explains and helps us understand many of the stylized facts discussed above. More important, the theory lets us predict how changes in economic conditions or in government policies will affect work incentives.

The representative person in our model receives satisfaction both from the consumption of goods (which we denote by C) and from the consumption of leisure (L). Obviously, the person consumes many different types of goods during any given period. To simplify matters, we aggregate the dollar value of all the goods that the person consumes and define C as the total dollar value of all the goods that the person purchases during the period. For example, if the person spends $1,000 weekly on food, rent, car payments, movie tickets, and other items, the variable C would take on the value of $1,000. The variable L gives the number of hours of leisure that a person consumes during the same time period.

Utility and Indifference Curves

The notion that individuals get satisfaction from consuming goods and leisure is summarized by the **utility function:**

$$U = f(C, L) \qquad \qquad \text{(2-5)}$$

The utility function transforms the person's consumption of goods and leisure into an index U that measures the individual's level of satisfaction or happiness. This index is called *utility.* The higher the level of index $U,$ the happier the person. We make the sensible assumption that buying more goods or having more leisure hours both increase the person's utility. In the jargon of economics, C and L are "goods," not "bads."

Suppose that a person is consuming $500 worth of consumption goods and 100 hours of leisure weekly (point Y in Figure 2-2). This particular consumption basket yields a particular level of utility to the person, say 25,000 utils. It is easy to imagine that different combinations of consumption goods and hours of leisure might yield the same level of utility. For example, the person might say that she would be indifferent to consuming $500 worth of goods and 100 hours of leisure or consuming $400 worth of goods and 125 hours of leisure. Figure 2-2 illustrates the many combinations of C and L that generate this particular level of utility. The locus of such points is called an **indifference curve**—and all points along this curve yield 25,000 utils.

FIGURE 2-2 Indifference Curves

Points X and Y lie on the same indifference curve and yield the same level of utility (25,000 utils); point Z lies on a higher indifference curve and yields more utility.
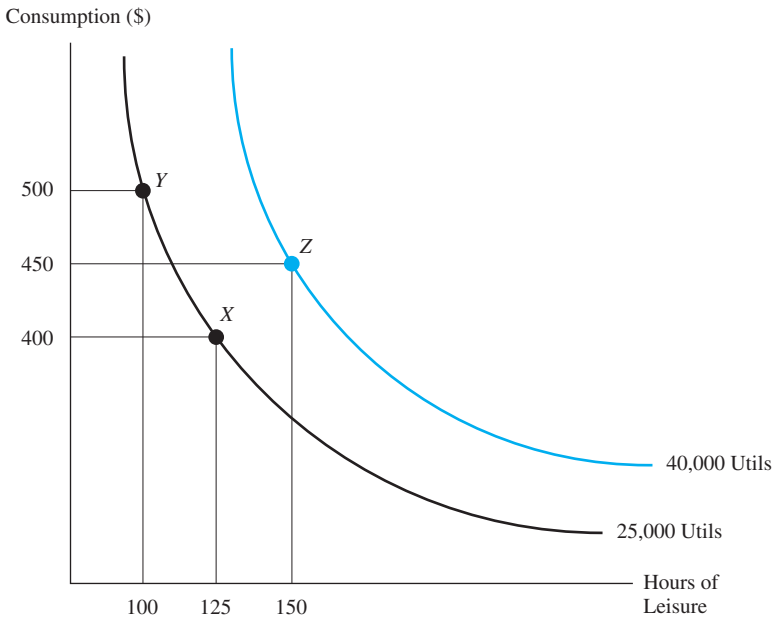
Consumption ($)

Suppose that the person were instead consuming $450 worth of goods and 150 hours of leisure (point Z in the figure). This consumption basket would put the person on a higher indifference curve, yielding 40,000 utils. We can then construct an indifference curve for this level of utility. In fact, we can construct an indifference curve for every level of utility. As a result, the utility function can be represented graphically in terms of a family (or a "map") of indifference curves.

Indifference curves have four important properties:

1. *Indifference curves are downward sloping.* We assumed that individuals prefer more of both C and L. If indifference curves were upward sloping, a consumption basket with more C and more L would yield the same level of utility as a consumption basket with less C and less L. This clearly contradicts our assumption that the individual likes both goods and leisure. The only way that we can offer a person a few more hours of leisure, and still hold utility constant, is to take away some of the goods.

2. *Higher indifference curves indicate higher levels of utility.* The consumption bundles lying on the indifference curve that yields 40,000 utils are preferred to the bundles lying on the curve that yields 25,000 utils. To see this, note that point Z in the figure must yield more utility than point X, simply because the bundle at point Z allows the person to consume more goods and leisure.

3. *Indifference curves do not intersect.* To see why, consider Figure 2-3, where indifference curves are allowed to intersect. Because points X and Y lie on the same indifference curve, the individual would be indifferent between the bundles X and Y. Because points Y and Z

FIGURE 2-3   **Indifference Curves Do Not Intersect**

Points *X* and *Y* yield the same utility because they are on the same indifference curve; points *Y* and *Z* also should yield the same utility. Point *Z,* however, is clearly preferable to point *X.*

Consumption ($)

Hours of
Leisure
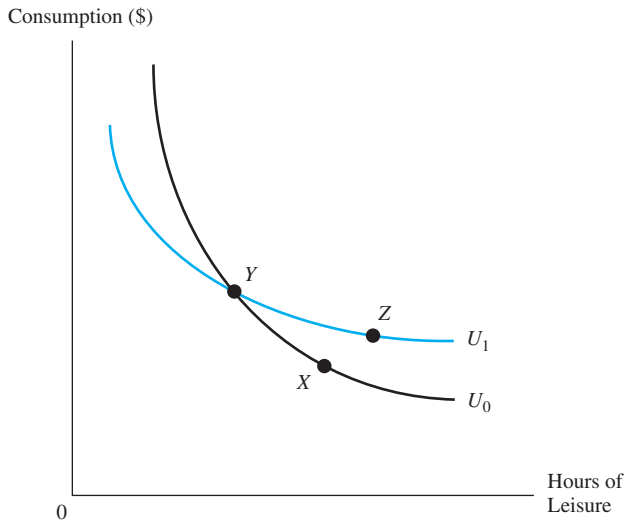
lie on the same indifference curve, the individual would be indifferent between bundles *Y* and *Z.* The person would then be indifferent between *X* and *Y,* and between *Y* and *Z,* so that she should also be indifferent between *X* and *Z.* But *Z* is clearly preferable to *X,* because *Z* has more goods and more leisure. Indifference curves that intersect contradict our assumption that individuals like to consume both goods and leisure.

4. *Indifference curves are convex to the origin.* The convexity of indifference curves does not follow from either the definition of indifference curves or the assumption that both goods and leisure are "goods." The convexity reflects an additional assumption about the shape of the utility function. It turns out (see problem 1 at the end of the chapter) that indifference curves must be convex to the origin if we are ever to observe a person sharing her time between work and leisure activities.

The Slope of an Indifference Curve

What happens to a person's utility as she allocates one more hour to leisure or buys an additional dollar's worth of goods? The **marginal utility** of leisure is defined as the change in utility resulting from an additional hour devoted to leisure activities, holding constant the amount of goods consumed. We denote the marginal utility of leisure as MU_L. Similarly, we can define the marginal utility of consumption as the change in utility if the individual consumes one more dollar's worth of goods, holding constant the number of hours devoted to leisure activities. We denote the marginal utility of consumption by MU_C. Because we have assumed that both leisure and the consumption of goods are desirable activities, the marginal utilities of leisure and consumption must be positive numbers.

As we move along an indifference curve, say from point *X* to point *Y* in Figure 2-2, the slope of the indifference curve measures the rate at which a person is willing to give up some leisure time in return for additional consumption, *while holding utility constant.* Put

differently, the slope tells us how many additional dollars' worth of goods it would take to "bribe" the person into giving up some leisure time. It can be shown that the slope of an indifference curve equals[7]

$$\frac{\Delta C}{\Delta L} = -\frac{MU_L}{MU_C} \qquad \text{(2-6)}$$

The absolute value of the slope of an indifference curve, which is also called the **marginal rate of substitution (MRS) in consumption,** is the ratio of marginal utilities.

The assumption that indifference curves are convex to the origin is essentially an assumption about how the marginal rate of substitution changes as the person moves along an indifference curve. Convexity implies that the slope of an indifference curve is steeper when the worker is consuming a lot of goods and little leisure, and that the curve is flatter when the worker is consuming few goods and a lot of leisure. As a result, the absolute value of the slope of an indifference curve declines as the person "rolls down" the curve. The assumption of convexity, therefore, is equivalent to an assumption of *diminishing* marginal rate of substitution.

Differences in Preferences across Workers

The map of indifference curves presented in Figure 2-2 illustrates the way a *particular* worker views the trade-off between leisure and consumption. Different workers will typically view this trade-off differently. In other words, some persons may like to devote a great deal of time and effort to their jobs, whereas other persons would prefer to devote most of their time to leisure. These interpersonal differences in preferences imply that the indifference curves may look quite different for different workers.

Figure 2-4 shows the indifference curves for two workers, Cindy and Mindy. Cindy's indifference curves tend to be very steep, indicating that her marginal rate of substitution takes on a very high value (see Figure 2-4*a*). In other words, she requires a sizable monetary bribe (in terms of additional consumption) to convince her to give up an additional hour of leisure. Cindy obviously likes leisure, and she likes it a lot. Mindy, on the other hand, has flatter indifference curves, indicating that her marginal rate of substitution takes on a low value (see Figure 2-4*b*). Mindy, therefore, does not require a large bribe to convince her to give up an additional hour of leisure.

Interpersonal differences in the "tastes for work" are obviously important determinants of differences in labor supply in the population. Workers who like leisure a lot (like Cindy) will tend to work few hours. And workers who do not attach a high value to their leisure time (like Mindy) will tend to be workaholics.

---

[7] To show that the slope of an indifference curve equals the ratio of marginal utilities, suppose that points X and Y in Figure 2-2 are very close to each other. In going from point X to point Y, the person is giving up ΔL hours of leisure, and each hour of leisure she gives up has a marginal utility of MU_L. Therefore, the loss in utility associated with moving from X to Y is given by $\Delta L \times MU_L$. The move from X to Y also involves a gain in utility. After all, the worker is not just giving up leisure time; she is consuming an additional ΔC dollars' worth of goods. Each additional dollar of consumption increases utility by MU_C units. The total gain in utility is given by $\Delta C \times MU_C$. By definition, all points along an indifference curve yield the same utility. This implies that the loss in moving from point X to point Y must be exactly offset by the gain, or $(\Delta L \times MU_L) + (\Delta C \times MU_C) = 0$. Equation (2-6) is obtained by rearranging terms.

FIGURE 2-4 **Differences in Preferences across Workers**

(*a*) Cindy's indifference curves are relatively steep, indicating that she requires a substantial bribe to give up an additional hour of leisure. (*b*) Mindy's indifference curves are relatively flat, indicating that she attaches a much lower value to her leisure time.
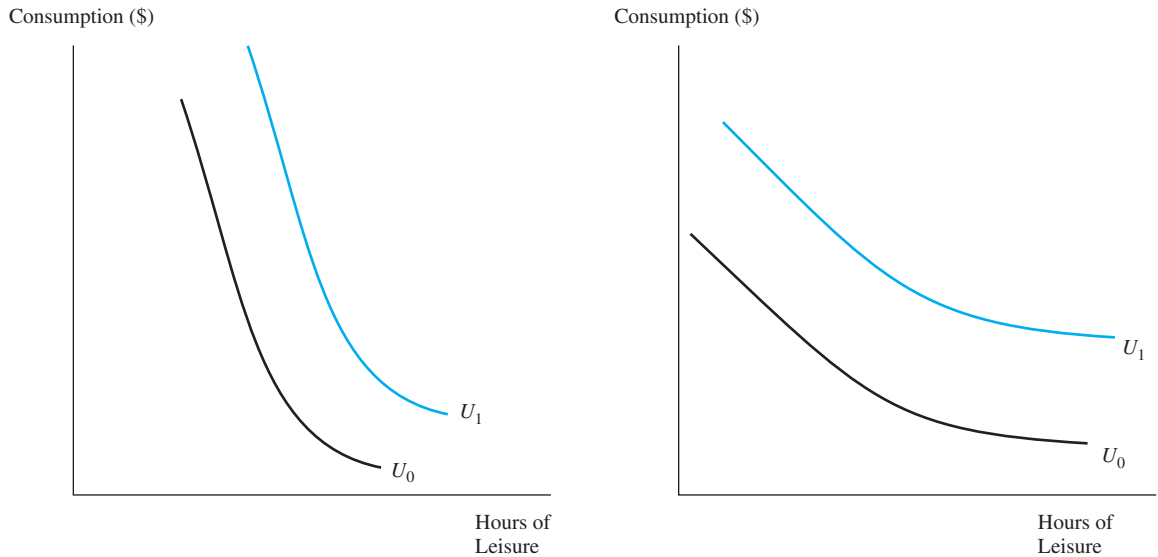
(*a*) Cindy's Indifference Curves

(*b*) Mindy's Indifference Curves

For the most part, economic models gloss over these interpersonal differences in preferences. The reason for this omission is that differences in tastes, although probably very important, are hard to observe and measure. It would be extremely difficult, if not impossible, to conduct surveys that would attempt to measure differences in indifference curves across workers. Moreover, the reliance on interpersonal differences in tastes provides an easy way out for anyone who wishes to explain why different workers behave differently. After all, one could simply argue that different behavior patterns between any two workers arise because worker *A* likes leisure more than worker *B,* and there would be no way of proving whether such a statement is correct or not.

Economic models instead stress the impact of variables that are much more easily observable—such as wages and incomes—on the labor supply decision. Because these variables can be observed and measured, the predictions made by the model about which types of persons will tend to work more are testable and refutable.

2-4   The Budget Constraint

The person's consumption of goods and leisure is constrained by her time and by her income. Part of the person's income (such as property income, dividends, and lottery prizes) is independent of how many hours she works. We denote this "nonlabor income" by *V.* Let *h* be the number of hours the person will allocate to the labor market during

the period and w be the hourly wage rate. The person's **budget constraint** can be written as

$$C = wh + V \qquad\qquad (2\text{-}7)$$

In words, the dollar value of expenditures on goods (C) must equal the sum of labor earnings (wh) and nonlabor income (V).[8]

As we will see, the wage rate plays a central role in the labor supply decision. Initially, we assume that the wage rate is constant *for a particular person,* so the person receives the same hourly wage regardless of how many hours she works. In fact, the "marginal" wage rate (that is, the wage rate received for the last hour worked) generally depends on how many hours a person works. Persons who work over 40 hours per week typically receive an overtime premium, and the wage rate in part-time jobs is often lower than the wage rate in full-time jobs.[9] For now, we ignore the possibility that a worker's marginal wage may depend on how many hours she chooses to work.

Given the assumption of a constant wage rate, it is easy to graph the budget constraint. The person has two alternative uses for her time: work or leisure. The total time allocated to each of these activities must equal the total time available in the period, say T hours per week, so that $T = h + L$. We can then rewrite the budget constraint as

$$C = w(T - L) + V \qquad\qquad (2\text{-}8)$$

or

$$C = (wT + V) - wL$$

This last equation is in the form of a line, and the slope is the negative of the wage rate (or $-w$).[10] The **budget line** is illustrated in Figure 2-5. Point E in the graph indicates that if the person decides not to work at all and devotes T hours to leisure activities, she can still purchase V dollars' worth of consumption goods. Point E is the *endowment point.* If the person is willing to give up one hour of leisure, she can then move up the budget line and purchase an additional w dollars' worth of goods. In fact, each additional hour of leisure that the person is willing to give up allows her to buy an additional w dollars' worth of goods. In other words, each hour of leisure consumed has a price, and the price is given by the wage rate. If the worker gives up all her leisure activities, she ends up at the intercept of the budget line and can buy $(wT + V)$ dollars' worth of goods.

The consumption and leisure bundles that lie below the budget line are available to the worker; the bundles that lie above the budget line are not. The budget line, therefore, delineates the frontier of the worker's **opportunity set**—the set of all the consumption baskets that a particular worker can afford to buy.

---

[8] The specification of the budget constraint implies that the worker does not save in this model. The worker spends all of her income in the period under analysis.

[9] Shelly Lundberg, "Tied Wage-Hours Offers and the Endogeneity of Wages," *Review of Economics and Statistics* 67 (August 1985): 405–410. There are also jobs, such as volunteer work, where the observed wage rate is zero; see Richard B. Freeman, "Working for Nothing: The Supply of Volunteer Labor," *Journal of Labor Economics* 15 (January 1997): S140–S166.

[10] Recall that the equation for a line relating the variables y and x is given by $y = a + bx$, where a is the intercept and b is the slope.

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 enter the labor market. The worker moves up the budget line as she trades off an hour of leisure for additional consumption. The absolute value of the slope of the budget line is the wage rate.
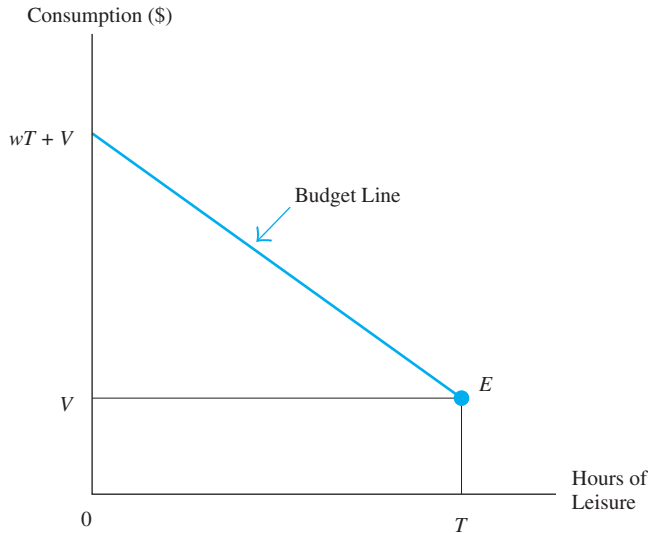
2-5   The Hours of Work Decision

We make one important assumption about the person's behavior: she wishes to choose the particular combination of goods and leisure that maximizes her utility. This means that the person will choose the level of goods and leisure that lead to the highest possible level of the utility index U—given the limitations imposed by the budget constraint.

Figure 2-6 illustrates the solution to this problem. As drawn, the budget line FE describes the opportunities available to a worker who has $100 of nonlabor income per week, faces a market wage rate of $10 per hour, and has 110 hours of nonsleeping time to allocate between work and leisure activities (assuming she sleeps roughly 8 hours per day).

Point P gives the optimal bundle of consumption goods and hours of leisure chosen by the utility-maximizing worker. The highest indifference curve attainable places her at point P and gives her U^* units of utility. At this solution, the worker consumes 70 hours of leisure per week, works a 40-hour workweek, and buys $500 worth of goods weekly. The worker would obviously prefer to consume a bundle on indifference curve U_1, which provides a higher level of utility. For example, the worker would prefer to be at point Y, where she works a 40-hour workweek and can purchase $1,100 worth of consumption goods. Given her wage and nonlabor income, however, the worker could never afford this consumption bundle. In contrast, the worker could choose a point such as A, which lies on the budget line, but she would not do so. After all, point A gives her less utility than point P.

The optimal consumption of goods and leisure for the worker, therefore, is given by the point where the budget line is tangent to the indifference curve. This type of solution is called an *interior solution* because the worker is not at either corner of the opportunity set (that is, at point F, working all available hours, or at point E, working no hours whatsoever).

FIGURE 2-6   An Interior Solution to the Labor-Leisure Decision

A utility-maximizing worker chooses the consumption-leisure bundle given by point P, where the indifference curve is tangent to the budget line.
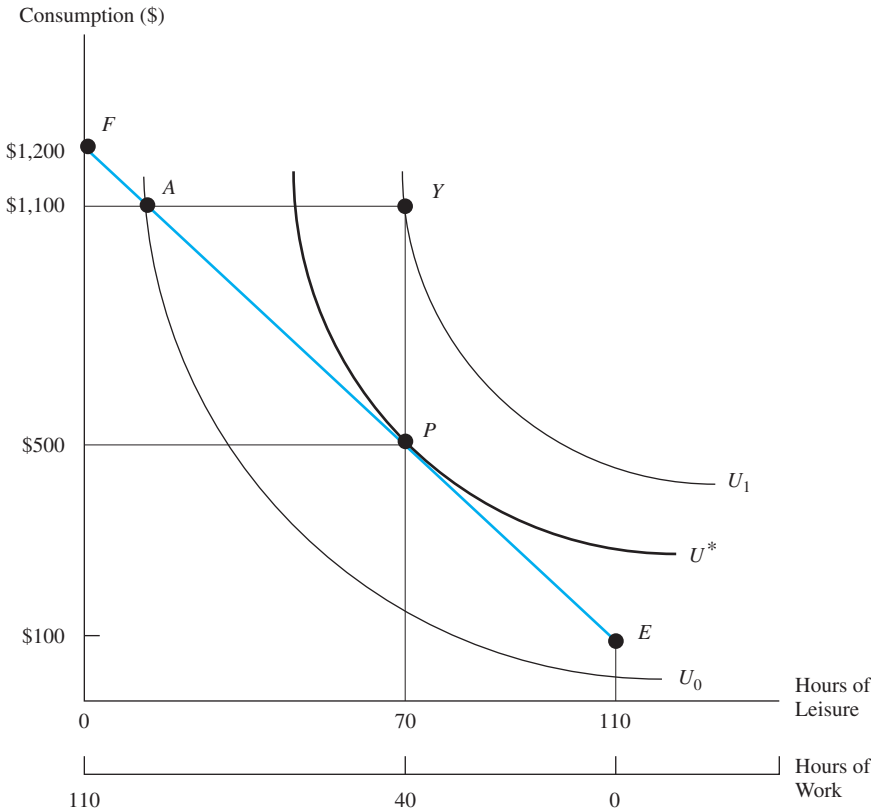
Interpreting the Tangency Condition

At the optimal point P, the budget line is tangent to the indifference curve. In other words, the slope of the indifference curve equals the slope of the budget line. This implies that[11]

$$\frac{MU_L}{MU_C} = w \qquad (2\text{-}9)$$

At the chosen level of consumption and leisure, the marginal rate of substitution (the rate at which a person is willing to give up leisure hours in exchange for additional consumption) equals the wage rate (the rate at which the market allows the worker to substitute one hour of leisure time for consumption).

[11] Although the slope of the indifference curve and the slope of the budget line are both negative numbers, the minus signs cancel out when the two numbers are set equal to each other, resulting in the condition reported in equation (2-9).

The economic intuition behind this condition is easier to grasp if we rewrite it as

$$\frac{MU_L}{w} = MU_C \qquad \text{(2-10)}$$

The quantity MU_L gives the additional utility received from consuming an extra hour of leisure. This extra hour costs w dollars. The left-hand side of equation (2-10), therefore, gives the number of utils received from spending an additional dollar on leisure. Because C is defined as the dollar value of expenditures on consumption goods, MU_C gives the number of utils received from spending an additional dollar on consumption goods. The tangency solution at point P in Figure 2-6 implies that the last dollar spent on leisure activities buys the same number of utils as the last dollar spent on consumption goods. If this equality did not hold (so that, for example, the last dollar spent on consumption buys more utils than the last dollar spent on leisure), the worker would not be maximizing utility. She could rearrange her consumption plan so as to purchase more of the commodity that yields more utility for the last dollar.

What Happens to Hours of Work When Nonlabor Income Changes?

We wish to determine what happens to hours of work when the worker's nonlabor income V increases. The increase in V might be triggered by the payment of higher dividends on the worker's stock portfolio or perhaps because some distant relatives had named the worker as the beneficiary in their will.

Figure 2-7 illustrates what happens to hours of work when the worker has an increase in V, *holding the wage constant*.[12] Initially, the worker's nonlabor income equals $100 weekly, which is associated with endowment point E_0. Given the worker's wage rate, the budget line is then given by F_0E_0. The worker maximizes utility by choosing the bundle at point P_0. At this point, the worker consumes 70 hours of leisure and works 40 hours.

The increase in nonlabor income to $200 weekly shifts the endowment point to E_1, so that the new budget line is given by F_1E_1. Because the worker's wage rate is being held constant, the slope of the budget line originating at point E_1 is the same as the slope of the budget line that originated at point E_0. An increase in nonlabor income that holds the wage constant expands the worker's opportunity set through a parallel shift in the budget line.

The increase in nonlabor income allows the worker to jump to a higher indifference curve, such as point P_1 in Figure 2-7. Increases in nonlabor income necessarily make the worker better off. After all, the expansion of the opportunity set opens up many additional opportunities for the worker. Figure 2-7*a* draws point P_1 so that the additional nonlabor income increases both expenditures on consumption goods and the number of leisure hours consumed. As a result, the length of the workweek falls to 30 hours. Figure 2-7*b* draws point P_1 so that the additional nonlabor income reduces the demand for leisure hours, increasing the length of the workweek to 50 hours. The impact of the change in nonlabor income (holding wages constant) on the number of hours worked is called an **income effect.**

---

[12] This type of theoretical exercise is called *comparative statics,* and is one of the main tools of economic theory. The methodology isolates how the outcomes experienced by a particular individual respond to a change in the value of one of the model's parameters. In this subsection, we are using the methodology to predict what should happen to labor supply when the worker's nonlabor income increases.

FIGURE 2-7 The Effect of a Change in Nonlabor Income on Hours of Work

An increase in nonlabor income leads to a parallel, upward shift in the budget line, moving the worker from point P_0 to point P_1. (a) If leisure is a normal good, hours of work fall. (b) If leisure is an inferior good, hours of work increase.
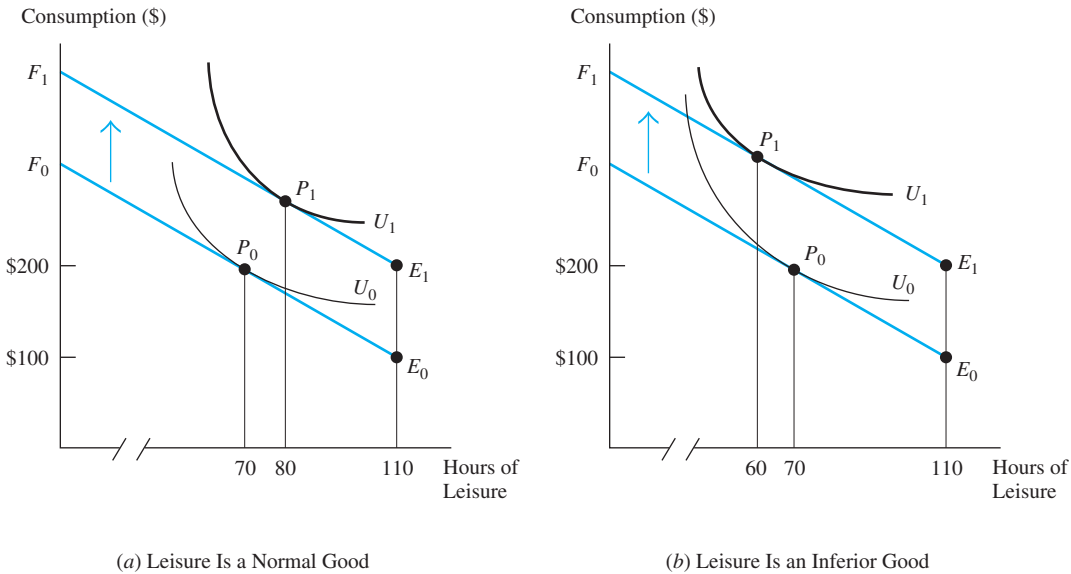
(a) Leisure Is a Normal Good   (b) Leisure Is an Inferior Good

Both panels in Figure 2-7 draw "legal" indifference curves. Both panels have indifference curves that are downward sloping, do not intersect, and are convex to the origin. It seems, therefore, that we cannot predict how an increase in nonlabor income affects hours of work unless we make an additional restriction on the shape of indifference curves. The additional restriction we make is that leisure is a "normal" good (as opposed to leisure being an "inferior" good).

We define a commodity to be a normal good when increases in income, holding the prices of all goods constant, increase its consumption. A commodity is an inferior good when increases in income, holding prices constant, decrease its consumption. Low-priced subcompact cars, such as the ill-fated Yugo, for instance, are typically thought of as inferior goods, whereas BMWs are typically thought of as normal goods. In other words, we would expect the demand for Yugos to decline as nonlabor income increased, and the demand for BMWs to increase.

If we reflect on whether leisure is a normal or an inferior good, most of us would probably reach the conclusion that leisure activities are a normal good. Put differently, if we were wealthier, we would surely demand a lot more leisure. We could then visit Aspen in December, Rio in February, and exotic beaches in the South Pacific in the summer.

Because it seems reasonable to assume that leisure is a normal good and because there is some evidence (discussed below) supporting this assumption, our discussion will focus on this case. The assumption that leisure is a normal good resolves the conflict between the two panels in Figure 2-7 in favor of the panel on the left-hand side. An increase in V then raises the demand for leisure hours and thus reduces hours of work. *The income effect, therefore, implies that an increase in nonlabor income, holding the wage rate constant, reduces hours of work.*

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.

(a)   (b)
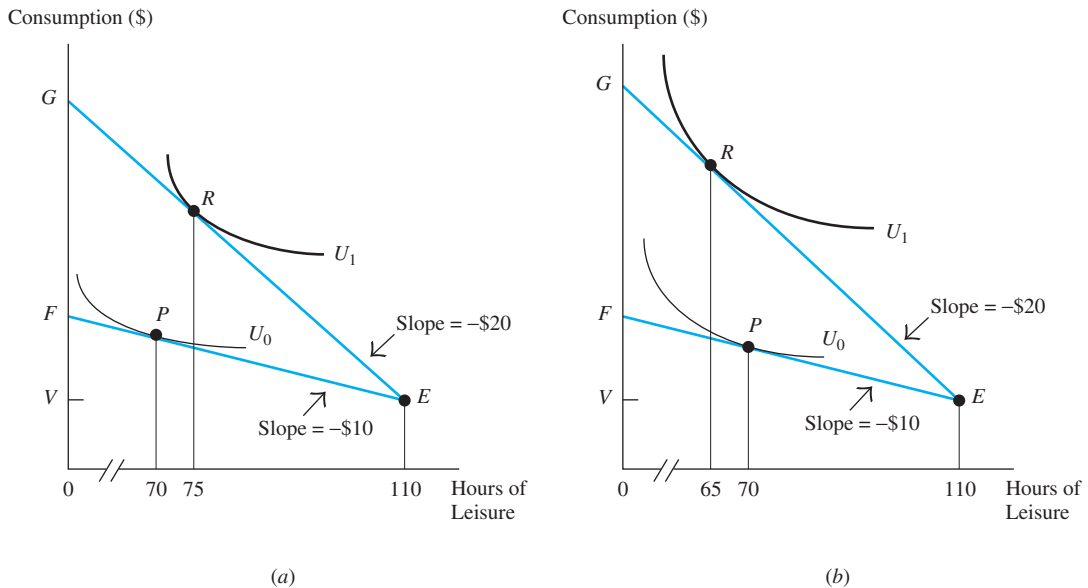
What Happens to Hours of Work When the Wage Changes?

Consider a wage increase from $10 to $20 an hour, holding nonlabor income *V* constant. The wage increase rotates the budget line around the endowment point, as illustrated in Figure 2-8. The rotation of the budget line shifts the opportunity set from *FE* to *GE*. It should be obvious that a wage increase does not change the endowment point: the dollar value of the goods that can be consumed when one does not work is the same regardless of whether the wage rate is $10 or $20 an hour.

The two panels presented in Figure 2-8 illustrate the possible effects of a wage increase on hours of work. In Figure 2-8*a,* the wage increase shifts the optimal consumption bundle from point *P* to point *R.* At the new equilibrium, the individual consumes more leisure (the increase is from 70 to 75 hours), so that hours of work fall from 40 to 35 hours.

Figure 2-8*b,* however, illustrates the opposite result. The wage increase again moves the worker to a higher indifference curve and shifts the optimal consumption bundle from point *P* to point *R.* This time, however, the wage increase reduces leisure hours (from 70 to 65 hours), so the length of the workweek increases from 40 to 45 hours. It seems, therefore, that we cannot make an unambiguous prediction about an important question without making even more assumptions.

The reason for the ambiguity in the relation between hours of work and the wage rate is of fundamental importance and introduces a set of tools and ideas that play a central role in all of economics. Both panels in Figure 2-8 show that, regardless of what happens to hours of work, a wage increase expands the worker's opportunity set. Put differently, a worker has more opportunities when she makes $20 an hour than when she makes $10 an hour. We know that an increase in income increases the demand for all normal goods, including leisure. The increase in the wage thus increases the demand for leisure, which reduces hours of work.

But this is not all that happens. The wage increase also makes leisure more expensive. When the worker earns $20 an hour, she gives up $20 every time she decides to take an hour off. As a result, leisure time is a very expensive commodity for high-wage workers and a relatively cheap commodity for low-wage workers. High-wage workers should then have strong incentives to cut back on their consumption of leisure activities. A wage increase thus reduces the demand for leisure and increases hours of work.

This discussion highlights the essential reason for the ambiguity in the relation between hours of work and the wage rate. A high-wage worker wants to enjoy the rewards of her high income, and hence would like to consume more leisure. The same worker, however, finds that leisure is very expensive and that she simply cannot afford to take time off from work.

These two conflicting forces are illustrated in Figure 2-9a. As before, the initial wage rate is $10 per hour. The worker maximizes her utility by choosing the consumption bundle given by point P, where she is consuming 70 hours of leisure and works 40 hours per week. Suppose the wage increases to $20. As we have seen, the budget line rotates and the new consumption bundle is given by point R. The worker is now consuming 75 hours of leisure and working 35 hours. As drawn, the person is working fewer hours at the higher wage.

It helps to think of the move from point P to point R as a two-stage move. The two stages correspond exactly to our discussion that the wage increase generates two effects: It increases the worker's income and it raises the price of leisure. To isolate the income effect, suppose we draw a budget line that is parallel to the old budget line (so that its slope is also –$10), but tangent to the new indifference curve. This budget line (DD) is also illustrated in Figure 2-9a, and generates a new tangency point Q.

FIGURE 2-9   Decomposing the Impact of a Wage Change into Income and Substitution Effects
An increase in the wage rate generates both income and substitution effects. The income effect (the move from point P to point Q) reduces hours of work; the substitution effect (the move from Q to R) increases hours of work.
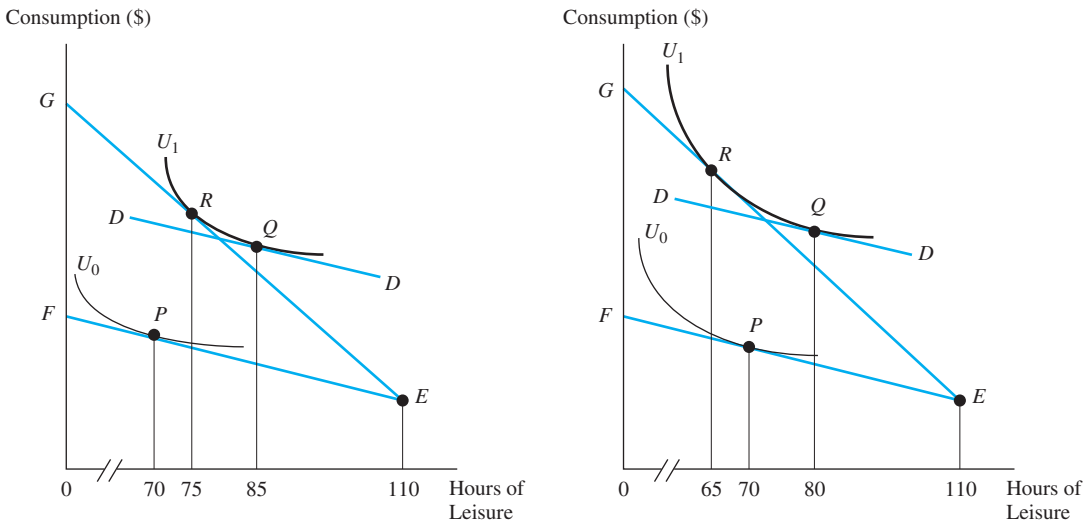
(a) Income Effect Dominates   (b) Substitution Effect Dominates

The move from initial position *P* to final position *R* can then be decomposed into a first-stage move from *P* to *Q* and a second-stage move from *Q* to *R*. It is easy to see that the move from point *P* to point *Q* is an income effect. In particular, the move from *P* to *Q* arises from a change in the worker's income, holding wages constant. The income effect isolates the change in the consumption bundle induced by the additional income generated by the wage increase. Because both leisure and goods are normal goods, point *Q* must lie to the northeast of point *P* (so that more is consumed of both goods and leisure). The income effect thus increases the demand for leisure (from 70 to 85 hours) and reduces hours of work by 15 hours per week.

The second-stage move from *Q* to *R* is called the **substitution effect.** It illustrates what happens to the worker's consumption bundle as the wage increases, holding utility constant. By moving along an indifference curve, the worker's utility or "real income" is held fixed. The substitution effect thus isolates the impact of the increase in the price of leisure on hours of work, holding real income constant.

The move from point *Q* to point *R* illustrates a substitution away from leisure time and toward consumption of other goods. In other words, as the wage rises, the worker devotes less time to expensive leisure activities (from 85 to 75 hours) and increases her consumption of goods. Through the substitution effect, therefore, the wage increase reduces the demand for leisure and increases hours of work by 10 hours. *The substitution effect implies that an increase in the wage rate, holding real income constant, increases hours of work.*

As drawn in Figure 2-9*a,* the decrease in hours of work generated by the income effect (15 hours) exceeds the increase in hours of work associated with the substitution effect (10 hours). The stronger income effect thus leads to a negative relationship between hours of work and the wage rate. In Figure 2-9*b,* the income effect (again the move from point *P* to point *Q*) decreases hours of work by 10 hours, whereas the substitution effect (the move from *Q* to *R*) increases hours of work by 15 hours. Because the substitution effect dominates, there is a positive relationship between hours of work and the wage rate.

The reason for the ambiguity in the relationship between hours of work and the wage rate should now be clear. As the wage rises, a worker faces a larger opportunity set and the income effect increases her demand for leisure and decreases labor supply. As the wage rises, however, leisure becomes more expensive and the substitution effect generates incentives for that worker to switch away from the consumption of leisure and instead consume more goods. This shift frees up leisure hours and thus increases hours of work.

To summarize the relation between hours of work and the wage rate:

- An increase in the wage rate increases hours of work if the substitution effect dominates the income effect.
- An increase in the wage rate decreases hours of work if the income effect dominates the substitution effect.

2-6   To Work or Not to Work?

Our analysis of the relation between nonlabor income, the wage rate, and hours of work assumed that the person worked both before and after the change in nonlabor income or the wage. Hours of work then adjusted to the change in the opportunity set. But what factors motivate a person to enter the labor force in the first place?

The implication that our demand for leisure time responds to its price is not very surprising. When the wage rate is high, we will find ways of minimizing the use of our valuable time, such as contact a ticket broker and pay very high prices for concert and theater tickets, rather than stand in line for hours to buy a ticket at face value. We will often hire a nanny or send our children to day care, rather than withdraw from the labor market. And we will consume many pre-prepared meals and order pizza or take-out Chinese food, rather than engage in lengthy meal preparations.

It turns out that our allocation of time responds to economic incentives even when there are no easy substitutes available, such as when we decide how many hours to sleep. Sleeping takes a bigger chunk of our time than any other activity, including market work. The typical man sleeps 56.0 hours per week, whereas the typical woman sleeps 56.9 hours per week. Although most persons think that how long we sleep is biologically (and

perhaps even culturally) determined, recent research suggests that, to some extent, sleep time also can be viewed as simply another activity that responds to economic incentives. As long as some minimum biological threshold for the length of a sleeping spell is met, the demand for sleep time seems to respond to changes in the price of time.

In particular, there is a negative correlation between a person's earnings capacity and the number of hours spent sleeping. More highly educated persons, for example, sleep less—an additional four years of school decreases sleep time by about one hour per week. Similarly, a 20 percent wage increase reduces sleep time by 1 percent, or about 34 minutes per week. When the wage is high, therefore, even dreaming of a nice, long vacation in a remote island becomes expensive.

To illustrate the nature of the work decision, consider Figure 2-10. The figure draws the indifference curve that goes through the endowment point E. This indifference curve indicates that a person who does not work at all receives U_0 units of utility. The woman, however, can choose to enter the labor market and trade some of her leisure time for earnings that will allow her to buy consumption goods. The decision of whether to work or not boils down to a simple question: Are the "terms of trade"—the rate at which leisure can be traded for additional consumption—sufficiently attractive to bribe her into entering the labor market?

Suppose initially that the person's wage rate is given by w_{low} so that the woman faces budget line *GE* in Figure 2-10. No point on this budget line can give her more utility than U_0. At this low wage, the person's opportunities are quite meager. If the worker were to move from the endowment point E to any point on the budget line *GE,* she would be moving to a lower indifference curve and be worse off. For example, at point X the woman gets only U_G utils. At wage w_{low}, therefore, the woman chooses not to work.

In contrast, suppose that the wage rate was given by w_{high}, so that the woman faces budget line *HE.* It is easy to see that moving to any point on this steeper budget line would increase her utility. At point Y, the woman gets U_H utils. At the wage w_{high}, therefore, the woman is better off working.

In sum, Figure 2-10 indicates that the woman does not enter the labor market at low wage rates (such as w_{low}), but does enter the labor market at high wage rates (such as w_{high}). As we rotate the budget line from wage w_{low} to wage w_{high}, we will typically encounter a wage rate, call it $\widetilde{w}$, that makes her indifferent between working and not working. We call $\widetilde{w}$ the

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 is given by the slope of the indifference curve at the endowment point.
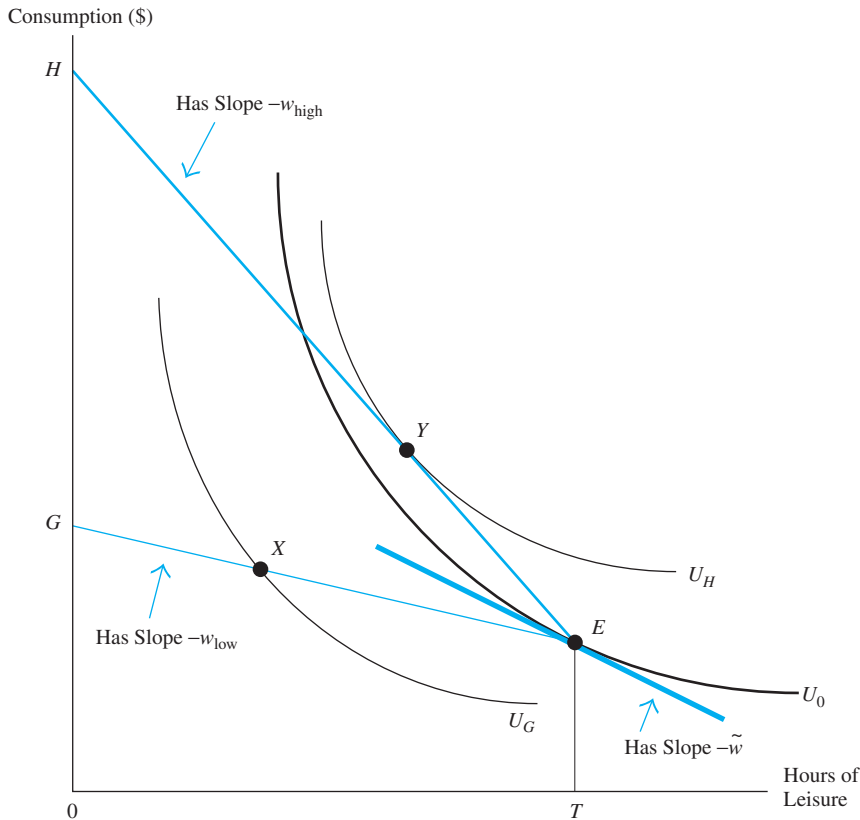
reservation wage. The reservation wage gives the minimum increase in income that would make a person indifferent between remaining at the endowment point E and working that first hour. In Figure 2-10, the reservation wage is given by the absolute value of the slope of the indifference curve at point E.

The definition of the reservation wage implies that the person will not work at all if the market wage is less than the reservation wage; and the person will enter the labor market if the market wage exceeds the reservation wage. The decision to work, therefore, is based on a comparison of the market wage, which indicates how much employers are willing to pay for an hour of work, and the reservation wage, which indicates how much the worker requires to be bribed into working that first hour.

The theory obviously implies that a high reservation wage makes it less likely that a person will work. The reservation wage will typically depend on the person's tastes for work, which helps to determine the slope of the indifference curve, as well on many other factors. For instance, the assumption that leisure is a normal good implies that the reservation wage

rises as nonlabor income increases.[13] Because workers want to consume more leisure as nonlabor income increases, a larger bribe will be required to convince a wealthier person to enter the labor market.

Holding the reservation wage constant, the theory also implies that high-wage persons are more likely to work. A rise in the wage rate, therefore, increases the labor force participation rate of a group of workers. As we shall see, this positive correlation between wage rates and labor force participation rates helps explain the rapid increase in the labor force participation rate of women observed in the United States and in many other countries in the past century.[14]

In sum, the theory predicts a positive relation between the person's wage rate and her probability of working. It is of interest to contrast this strong prediction with our earlier result that a wage increase has a theoretically ambiguous effect on hours of work, depending on whether the income or substitution effect dominates.

The disparity between these two results arises because an increase in the wage generates an income effect *only if the person is already working.* A person working 40 hours per week will surely be able to consume many more goods when the wage is $20 per hour than when the wage is $10 per hour. This type of wage increase makes leisure more expensive (so that the worker wants to work more) *and* makes the person wealthier (so that the worker wants to work less). In contrast, if the person is not working at all, an increase in the wage rate has no effect on her real income. The amount of goods that a nonworker can buy is independent of whether her potential wage rate is $10 or $20 an hour. An increase in the wage of a nonworker, therefore, does not generate an income effect. The wage increase simply makes leisure time more expensive and hence is likely to draw the nonworker into the labor force.

2-7   The Labor Supply Curve

The predicted relation between hours of work and the wage rate is called the **labor supply curve.** Figure 2-11 illustrates how a person's labor supply curve can be derived from the utility-maximization problem that we solved earlier.

The left panel of the figure shows the person's optimal consumption bundle at a number of alternative wage rates. As drawn, the wage of $10 is the person's reservation wage, the wage at which she is indifferent between working and not working. This person, therefore, supplies zero hours to the labor market at any wage less than or equal to $10. Once the wage rises above $10, the person chooses to work some hours. For example, she works

---

[13] Try to prove this statement by drawing a vertical line through the indifference curves in Figure 2-6. By moving up this vertical line, we are holding constant hours of leisure. Because of their convexity, the indifference curves will get steeper as we move to higher indifference curves.

[14] The modern analysis of labor force participation decisions within an economic framework began with the classic work of Jacob Mincer, "Labor Force Participation of Married Women," in H. Gregg Lewis, editor, *Aspects of Labor Economics,* Princeton, NJ: Princeton University Press, 1962, pp. 63–97. An important study that stresses the comparison between reservation and market wages is given by James J. Heckman, "Shadow Prices, Market Wages and Labor Supply," *Econometrica* 42 (July 1974): 679–694.

In 1970, there were only two state lotteries in the United States. These lotteries sold $100 million in tickets during the year. By 1996, 36 states participated in lotteries, and consumers purchased more than $34 billion of lottery tickets. The first prize in these lotteries sometimes reaches astronomical amounts. Consider, for example, the $314.9 million jackpot in the Powerball Lottery held on December 25, 2002, in 23 states, the District of Columbia, and the U.S. Virgin Islands. The demand for a chance at this fortune was so intense that long lines formed at many of the stores and shops that sold tickets. Pennsylvania lottery officials estimated that 600 tickets were being sold *per second* on Christmas Eve, despite the 1 in 120 million odds of winning the jackpot.

Thousands of players have become "instant millionaires" (although the payout is often stretched over a 20- or 30-year period). The Massachusetts official who hands out the initial checks to the winners reports that most new millionaires claim that the money will not change their lives. The neoclassical model of labor-leisure choice, however, predicts otherwise. Winning the lottery is a perfect example of an unexpected and often substantial increase in nonlabor income. As long as leisure is a normal good, we would predict that lottery winners would reduce their hours of work, and perhaps even withdraw entirely from the labor market.

An extensive study of the labor supply behavior of 1,000 lottery winners who won a jackpot of more than $50,000 is revealing. Nearly 25 percent of the winners (and of their spouses) left the labor force within a year, and an additional 9 percent reduced the number of hours they worked or quit a second job. Not surprisingly, the labor supply effects of lottery income depended on the size of the jackpot. Only 4 percent of the winners who won a jackpot between $50,000 and $200,000 left the labor force, but nearly 40 percent of those whose jackpot exceeded $1 million retired to the "easy life."

The experience of David Sneath, who worked at a Ford Motor Company warehouse for 34 years, says everything that needs to be said about income effects. After picking up his first payment on a $136 million jackpot, "I yelled to the boss, 'I'm out of here.'"

Sources: Roy Kaplan, "Lottery Winners and Work Commitment: A Behavioral Test of the American Work Ethic," *Journal of the Institute for Socioeconomic Studies* 10 (Summer 1985): 82–94; Charles T. Clotfelter and Philip J. Cook, *Selling Hope: State Lotteries in America,* Cambridge, MA: Harvard University Press, 1989; Guido W. Imbens, Donald B. Rubin, and Bruce Sacerdote, "Estimating the Effect of Unearned Income on Labor Supply, Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players," *American Economic Review* 91 (September 2001): 778–794; and www.msnbc.msn.com/id/23958892.

20 hours when the wage is $13; 40 hours when the wage if $20; and 30 hours when the wage is $25. Note that, as drawn, the figure implies that substitution effects dominate at lower wages and that income effects dominate at higher wages.

The right panel of the figure traces out the labor supply curve, the relation between the optimal number of hours worked and the wage rate. Initially, the labor supply curve is positively sloped as hours and wages move together. Once the wage rises above $20, however, the income effect dominates and hours of work decline as the wage rises, creating a segment of the labor supply curve that has a negative slope. The type of labor supply curve illustrated in Figure 2-11*b* is called a *backward-bending* labor supply curve because it eventually bends around and has a negative slope.

We can use the utility-maximization framework to derive a labor supply curve for every person in the economy. The labor supply curve in the aggregate labor market is then given by adding up the hours that all persons in the economy are willing to work at a given wage. Figure 2-12 illustrates how this "adding up" is done in an economy with two workers, Alice and Brenda. Alice has reservation wage $\tilde{w}_A$; Brenda has a higher reservation wage $\tilde{w}_B$. It should

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.
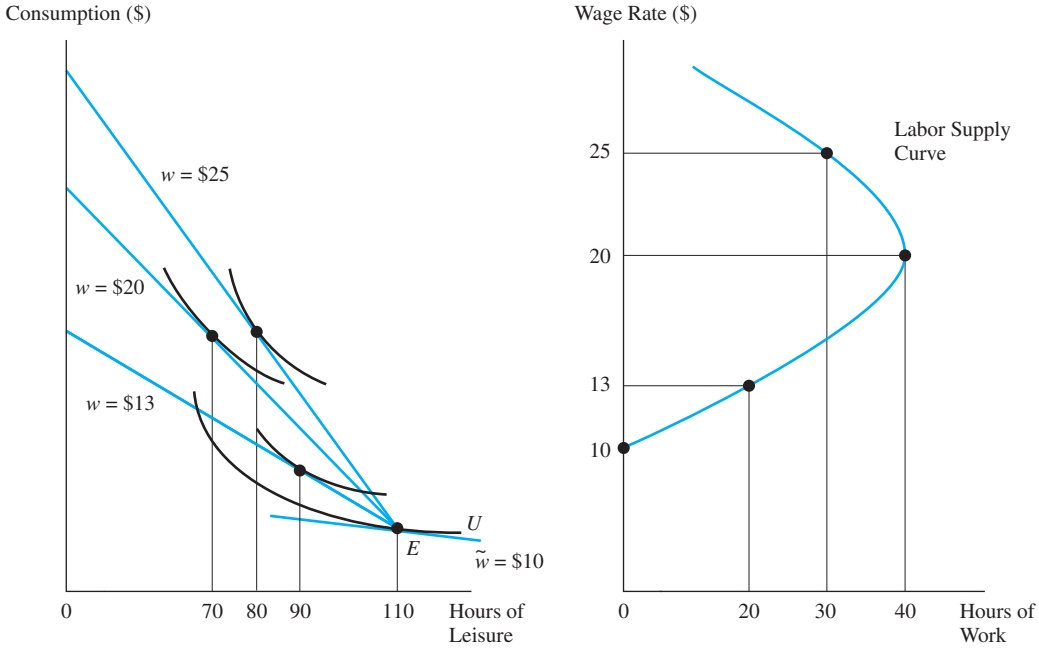
(a) Optimal Consumption Bundles

(b) Relation between Optimal Hours of Work and the Wage Rate

FIGURE 2-12   Derivation of the Market Labor Supply Curve from the Supply Curves of Individual Workers

The market labor supply curve "adds up" the supply curves of individual workers. When the wage is below $\tilde{w}_A$, no one works. As the wage rises, Alice enters the labor market. If the wage rises above $\tilde{w}_B$, Brenda enters the market.

(a) Alice

(b) Brenda

(c) Market

be clear that no one would work if the wage is below $\tilde{w}_A$, and that only Alice would work if the wage is between $\tilde{w}_A$ and $\tilde{w}_B$. At wages higher than $\tilde{w}_B$, market labor supply is given by the total number of hours worked by Alice and Brenda, or $h_A + h_B$. The labor supply curve in the market, therefore, is obtained by adding up the supply curves of all workers *horizontally*.

To measure the responsiveness of hours of work to changes in the wage rate, we define the **labor supply elasticity** as

$$\sigma = \frac{\text{Percent change in hours of work}}{\text{Percent change in wage rate}} = \frac{\Delta h / h}{\Delta w / w} = \frac{\Delta h}{\Delta w} \cdot \frac{w}{h} \quad \textbf{(2-11)}$$

The labor supply elasticity gives the percentage change in hours of work associated with a 1 percent change in the wage rate. The sign of the labor supply elasticity depends on whether the labor supply curve is upward sloping ($\Delta h / \Delta w$ (0) or downward sloping ($\Delta h / \Delta w < 0$), and, hence, is positive when substitution effects dominate and negative when income effects dominate. Hours of work are more responsive to changes in the wage the greater the absolute value of the labor supply elasticity.

To see how the labor supply elasticity is calculated, consider the following example. Suppose that the worker's wage is initially \$10 per hour and that she works 1,900 hours per year. The worker gets a raise to \$20 per hour, and she decides to work 2,090 hours per year. This worker's labor supply elasticity can then be calculated as

$$\sigma = \frac{\text{Percent change in hours of work}}{\text{Percent change in wage rate}} = \frac{10\%}{100\%} = 0.1 \quad \textbf{(2-12)}$$

When the labor supply elasticity is less than one in absolute value, the labor supply curve is said to be *inelastic.* In other words, there is relatively little change in hours of work for a given change in the wage rate. If the labor supply elasticity is greater than one in absolute value—indicating that hours of work are greatly affected by the change in the wage—the labor supply curve is said to be *elastic.* It is obvious that labor supply is inelastic in the numerical example in equation (2-12). After all, a doubling of the wage (a 100 percent increase) raised labor supply by only 10 percent.

2-8   Estimates of the Labor Supply Elasticity

Few topics in applied economics have been as thoroughly researched as the empirical relationship between hours of work and wages. We begin our review of this literature by focusing on the estimates of the labor supply elasticity for men. Since most prime-age men participate in the labor force, the typical study uses the sample of working men to correlate a particular person's hours of work with his wage rate and nonlabor income. In particular, the typical regression model estimated in these studies is

$$h_i = \beta w_i + \gamma V_i + \text{Other variables} \quad \textbf{(2-13)}$$

where h_i gives the number of hours that person i works; w_i gives his wage rate; and V_i gives his nonlabor income. The coefficient β measures the impact of a one-dollar wage increase on hours of work, holding nonlabor income constant; and the coefficient measures the impact of a one-dollar increase in nonlabor income, holding the wage constant. The neoclassical model of labor-leisure choice implies that the sign of the coefficient β depends on whether

income or substitution effects dominate. In particular, β is negative if income effects dominate and positive if substitution effects dominate. The estimate of the coefficient β can be used to calculate the labor supply elasticity defined by equation (2-11). Assuming leisure is a normal good, the theory also predicts that the coefficient γ should be negative because workers with more nonlabor income consume more leisure.

There are almost as many estimates of the labor supply elasticity as there are empirical studies in the literature. As a result, the variation in the estimates of the labor supply elasticity is enormous. Some studies report the elasticity to be zero; other studies report it to be large and negative; still others report it to be large and positive. There have been some attempts to determine which estimates are most credible.[15] These surveys conclude that the elasticity of the male labor supply is roughly around −0.1. In other words, a 10 percent increase in the wage leads, on average, to a 1 percent decrease in hours of work for men. In terms of the decomposition into income and substitution effects, there is some consensus that a 10 percent increase in the wage increases hours of work by about 1 percent because of the substitution effect, but also leads to a 2 percent decrease because of the income effect. As predicted by the theory, therefore, the substitution effect is positive.

Three key points are worth noting about the −0.1 "consensus" estimate of the labor supply elasticity. First, it is negative, so income effects dominate. The dominance of income effects is often used to explain the decline in hours of work between 1900 and 2000 that we documented earlier in this chapter. In other words, the secular decline in hours of work can be attributed to the income effects associated with rising real wages.[16] Second, the labor supply curve is inelastic. Hours of work for men do not seem to be very responsive to changes in the wage. In fact, one would not be stretching the truth too far if one were to claim that the male labor supply elasticity is essentially zero. This result should not be too surprising since most prime-age men work a full workweek every week of the year.[17]

---

[15] A recent survey of the labor supply literature is given by Richard Blundell and Thomas MaCurdy, "Labor Supply: A Review of Alternative Approaches," in Orley C. Ashenfelter and David Card, editors, *Handbook of Labor Economics,* vol. 3A, Amsterdam: Elsevier, 1999, pp. 1559–1695. Many of the large positive elasticities reported in the literature are found in studies that attempt to estimate the impact of changes in income tax rates on labor supply. A good survey of this literature is given by Jerry A. Hausman, "Taxes and Labor Supply," in Alan J. Auerbach and Martin Feldstein, editors, *Handbook of Public Economics,* vol. 1, Amsterdam: Elsevier, 1985, pp. 213–263. Recent research, however, suggests that a more careful specification of the econometric model used to estimate how taxes affect labor supply yields a labor supply response that is much weaker and closer in line with the consensus estimate of −0.1; see Thomas MaCurdy, David Green, and Harry Paarsch, "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply," *Journal of Human Resources* 25 (Summer 1990): 415–490; James P. Ziliak and Thomas J. Kniesner, "The Effect of Income Taxation on Consumption and Labor Supply," *Journal of Labor Economics* 23 (October 2005): 769–796.

[16] Thomas J. Kniesner, "The Full-Time Workweek in the United States: 1900–1970," *Industrial and Labor Relations Review* 30 (October 1976): 3–15; and John Pencavel, "A Cohort Analysis of the Association between Work Hours and Wages among Men," *Journal of Human Resources* 37 (Spring 2002): 251–274. In recent years, hours of work have begun to rise for highly educated men, high-wage men. This increase may be due to a strong substitution effect caused by a rapidly rising real wage; see Peter Kuhn and Fernando Luzano, "The Expanding Workweek? Understanding Trends in Long Work Hours among U.S. Men, 1979–2006," *Journal of Labor Economics* 26 (April 2008): 311–343.

[17] Recall, however, that the labor force participation rate of men fell throughout much of the twentieth century. For a study of this trend, see Chinhui Juhn, "The Decline of Male Labor Market Participation: The Role of Market Opportunities," *Quarterly Journal of Economics* 107 (February 1992): 79–121.

And, third, it is important to keep in mind that this is the "consensus" estimate of the labor supply elasticity for prime-age men. The available evidence suggests that the labor supply elasticity probably differs greatly between men and women and between younger and older workers.

Problems with the Estimated Elasticities

Why is there so much variation in the estimates of the labor supply elasticity across studies? It turns out that much of the empirical research in this area is marred by a number of statistical and measurement problems. In fact, each of the three variables that are crucial for estimating the labor supply model—the person's hours of work, the wage rate, and nonlabor income—introduces difficult problems into the estimation procedure.

Hours of Work

What precisely do we mean by hours of work when we estimate a labor supply model: is it hours of work per day, per week, or per year? The elaborate theoretical apparatus that we have developed does not tell us what the span of the time period should be. It turns out, however, that the observed responsiveness of hours of work to a wage change depends crucially on whether we look at a day, a week, or a year.[18] Not surprisingly, the labor supply curve becomes more elastic the longer the time period over which the hours-of-work variable is defined, so labor supply is almost completely inelastic if we analyze hours of work per week, but it is a bit more responsive if we analyze hours of work per year. Our conclusion that the labor supply elasticity is around -0.1 is based on studies that look at variation in annual hours of work.

There is also substantial measurement error in the hours-of-work measure that is typically reported in survey data.[19] Workers who are paid by the hour know quite well how many hours they worked last week; after all, their take-home pay depends directly on the length of the workweek. Many of us, however, are paid an annual salary and we make little (if any) effort to track exactly how many hours we work in any given week. When we are asked how many hours we work per week, many of us will respond "40 hours" because that is the easy answer. Actual hours of work, however, may have little to do with the mythical 40-hour workweek for many salaried workers. As we will see shortly, this measurement error introduces a bias into the estimation of the labor supply elasticity.

The Wage Rate

The typical salaried worker is paid an annual salary, regardless of how many hours she puts into her job. It is customary to define the wage rate of salaried workers in terms of the average wage, the ratio of annual earnings to annual hours worked. This calculation transmits any measurement errors in the reported measure of hours of work to the wage rate.

---

[18] See Finis Welch, "Wages and Participation," *Journal of Labor Economics* 15 (January 1997): S77–S103; and Chinhui Juhn, Kevin M. Murphy, and Robert H. Topel, "Why Has the Natural Rate of Unemployment Increased over Time?" *Brookings Papers on Economic Activity* 2 (1991): 75–126.

[19] John Bound, Charles Brown, Greg Duncan, and Willard Rogers, "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data," *Journal of Labor Economics* 12 (July 1994): 345–368.

To illustrate the problem introduced by these measurement errors, suppose that a worker overreports her hours of work. Because of the way the wage rate is constructed (that is, as the ratio of annual earnings to annual hours of work), the denominator of this ratio is too big and we estimate an artificially low wage rate. High reported hours of work are then associated with low wage rates, generating a spurious negative correlation between hours and average wages. Suppose instead that the worker underreports her hours of work. The constructed wage rate is then artificially high, again generating a spurious negative correlation between hours of work and the wage. As a result, measurement error tends to exaggerate the importance of income effects. In fact, there is evidence that correcting for measurement error in hours of work greatly reduces the magnitude of the income effect.[20]

Even in the absence of measurement error, there is an important conceptual problem in defining the wage rate as the ratio of annual earnings to hours of work for salaried workers.

[20] George J. Borjas, "The Relationship between Wages and Weekly Hours of Work: The Role of Division Bias," *Journal of Human Resources* 15 (Summer 1980): 409–423.

The correct price of leisure in the neoclassical model of labor-leisure choice is the marginal wage, the increase in earnings associated with an additional hour of work. The relevant marginal wage for salaried workers may have little to do with the average wage earned per hour.

Finally, a researcher attempting to estimate the labor supply model quickly encounters the serious problem that the wage rate is not observed for people who are not working. However, a person who is out of the labor market does *not* have a zero wage rate. All that we really know is that this person's wage is below the reservation wage. Many empirical studies avoid the problem of calculating the wages of nonworkers by simply throwing the nonworkers out of the sample that is used for calculating the labor supply elasticity.

This procedure, however, is fundamentally flawed. The decision of whether to work depends on a comparison of market wages and reservation wages. Persons who do not work have either very low wage rates or very high reservation wages. The sample of workers (or of nonworkers), therefore, is not a random sample of the population. Because most of the econometric techniques and statistical tests that have been developed specifically assume that the sample under analysis is a random sample, these techniques cannot be used to analyze the labor supply behavior of a sample that only includes workers. As a result, the estimated labor supply elasticities are not calculated correctly. This problem is typically referred to as "selection bias."[21]

Nonlabor Income

We would ideally like *V* to measure that part of the worker's income stream that has nothing to do with how many hours he works. For most people, however, the current level of nonlabor income partly represents the returns to past savings and investments. Suppose that some workers have a "taste for work." The shape of their indifference curves is such that they worked long hours, had high labor earnings, and were able to save and invest a large fraction of their income in the past. These are precisely the workers who will have high levels of nonlabor income today. If a worker's taste for work does not change over time, these are also the workers who will tend to work more hours today. The correlation between nonlabor income and hours of work will then be positive, simply because persons who have large levels of nonlabor income are the persons who tend to work many hours.

In fact, some studies in the literature report that workers who have more nonlabor income work more hours. This finding would suggest either that leisure is an inferior good or that the biases introduced by the correlation between tastes for work and nonlabor income are sufficiently strong to switch the sign of the estimated income effect. More careful studies that account for the correlation between "tastes for work" and nonlabor income find that increases in nonlabor income do indeed reduce hours of work.[22]

---

[21] A number of sophisticated statistical techniques have been developed to handle the self-selection problem. These techniques typically involve estimating labor supply functions that include not only the wage rate and nonlabor income as independent variables but also the predicted probability that a person is working. See James J. Heckman, "Sample Selection Bias as a Specification Error," *Econometrica* 47 (January 1979): 153–162; and James J. Heckman, "Sample Selection Bias as a Specification Error with an Application to the Estimation of Labor Supply Functions," in James P. Smith, editor, *Female Labor Supply: Theory and Estimation,* Princeton, NJ: Princeton University Press, 1980, pp. 206–248.

[22] James P. Smith, "Assets and Labor Supply," in Smith, editor, *Female Labor Supply: Theory and Estimation,* pp. 166–205.

2-9   Labor Supply of Women

Table 2-4 documents the growth of the female labor force in a number of countries between 1980 and 2003.[23] These statistics suggest two key results. There are substantial differences across countries in women's labor force participation rates. In Italy, for instance, fewer than half of women aged 15 to 64 participated in the labor force in 2003; in the United States and Canada, the participation rate hovered around 70 percent. These differences can probably be attributed to differences in economic variables and cultural factors, as well as the institutional framework in which labor supply decisions are being made.

Despite the international differences in the level of labor force participation, the data also reveal that these countries experienced a common trend: rising female labor force participation during the past few decades. The participation rate of women increased from 40 to 47 percent in Italy between 1980 and 2003; from 55 to 64 percent in Japan; and from 33 to 50 percent in Greece.

In the United States, the participation rate has grown over time both for a particular group of female workers and across cohorts of workers.[24] In other words, the participation rate of a given birth cohort of women increases as the women get older (past the childbearing years). For example, the participation rate of women born around

TABLE 2-4
International Differences in Female Labor Force Participation Rate (women aged 15–64)

Source: U.S. Bureau of the Census, *Statistical Abstract of the United States, 2006,* Washington, DC: Government Printing Office, Table 1343.

Country	1980	1990	2003
Australia	52.7	62.1	66.4
Canada	57.8	67.6	70.4
France	54.4	57.8	62.0
Germany	52.8	56.7	64.0
Greece	33.0	43.6	50.2
Ireland	36.3	43.8	56.2
Italy	39.6	45.9	46.8
Japan	54.8	60.3	64.2
Korea, South	—	51.2	54.3
Mexico	33.7	—	42.4
New Zealand	44.6	63.0	67.6
Portugal	54.3	62.9	67.2
Spain	32.2	41.2	50.7
Sweden	74.1	80.4	75.0
Turkey	—	36.7	26.9
United Kingdom	58.3	66.5	67.8
United States	59.7	68.5	71.7

[23] A survey of these international trends is given by Jacob Mincer, "Intercountry Comparisons of Labor Force Trends and of Related Developments: An Overview," *Journal of Labor Economics* 3 (January 1985 Supplement): S1–S32.

[24] James P. Smith and Michael P. Ward, "Time-Series Growth in the Female Labor Force," *Journal of Labor Economics* 3 (January 1985, Part 2): S59–S90; and Claudia Goldin, "Life-Cycle Labor-Force Participation of Married Women: Historical Evidence and Implications," *Journal of Labor Economics* 7 (January 1989): 20–47.

1930 was 27.7 percent when they were 30 years old and rose to 58.0 percent when they were 50 years old. Equally important, there has been a substantial increase in labor force participation across cohorts, with more recent cohorts having larger participation rates. At age 30, for example, women born around 1950 had a participation rate of 61.6 percent, more than twice the participation rate of women born in 1930 at an equivalent point in the life cycle.

Our theoretical discussion highlights the role of changes in the wage rate as a key determinant of the increase in female labor force participation. In particular, as the wage increases, nonworking women have an incentive to reduce the time they allocate to the household sector and are more likely to enter the labor market.[25] In fact, the real wage of women increased substantially in most countries. Between 1960 and 1980, the real wage of women grew at an annual rate of 6.2 percent for Australian women, 4.2 percent for British women, 5.6 percent for Italian women, and 2.1 percent for American women. The across-country relationship between the increase in labor force participation rates and the increase in the real wage is illustrated in Figure 2-13. Even without the use of sophisticated econometrics, one can see that labor force participation rates grew fastest in those developed countries that experienced the highest increase in the real wage.

FIGURE 2-13   **Cross-Country Relationship between Growth in Female Labor Force and the Wage, 1960–1980**

Source: Jacob Mincer, "Intercountry Comparisons of Labor Force Trends and of Related Developments: An Overview," *Journal of Labor Economics* 3 (January 1985, Part 2): S2, S6.
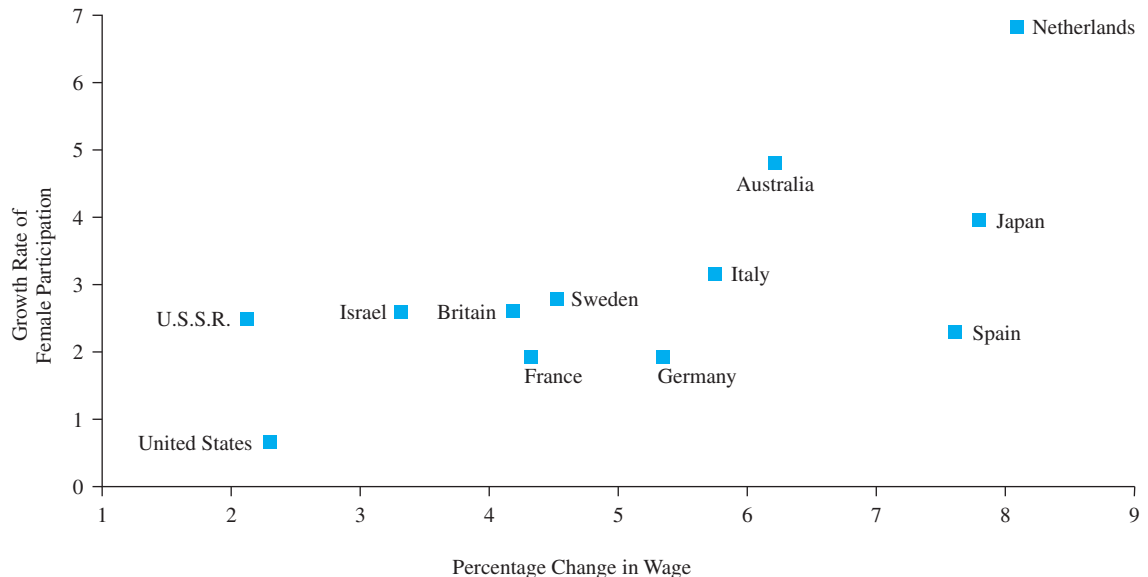
[25] Recall that the theory implies that a wage increase does not generate an income effect for non-workers. The only impact of a wage increase on this group of persons is to increase the price of leisure and to make it more likely that they will now enter the labor force.

The labor force participation decision is based on a comparison of the market wage with the reservation wage. Hence, the increase in the labor force participation rates of women could be due not only to a rise in the market wage but also to a decline in women's reservation wages. It is likely that an increase in the number of children raises a woman's reservation wage and reduces the probability that the woman will work. In fact, if a woman has children under the age of six, her probability of working falls by nearly 20 percentage points.[26] Between 1950 and 2000, the total lifetime fertility of the average adult woman declined from 3.3 to 2.1 children, so the reduction in fertility probably contributed to the increase in female labor force participation.[27] It is also likely, however, that the rise in the market wage, which increased female participation rates, also made childbearing a very expensive household activity. As a result, some of the causation runs in the opposite direction: women participate more not because they have fewer children; rather, they have fewer children because the rising wage induces them to reduce their time in the household sector and enter the labor market.[28]

More generally, the model suggests that women's labor supply may be more responsive to wage changes than men's labor supply. Note that a wage increase makes household production relatively less valuable at the same time that it increases the price of leisure. Therefore, a wage increase would encourage a person to substitute time away from household production and toward market work. A rise in the real wage will then draw many women out of the household production sector and move them into the market sector. Because very few men specialized in household production in earlier decades, such a transition would have been relatively rare among men.

Female labor force participation rates also are influenced by technological changes in the process of household production. There have been remarkable time-saving technological advances in household production, including stoves, washing machines, and the microwave oven. As a result, the amount of time required to produce many household commodities was cut drastically in the twentieth century, freeing up the scarce time for leisure activities and for work in the labor market. A large difference in the marginal product of household time between the husband and the wife makes it likely that one of the two spouses will specialize in the household sector. The technological advances in household production probably reduced the gap in household productivity between the two spouses, lessening the need for specialization and further contributing to the increase in female labor force participation rates.

The economic model should not be interpreted as saying that *only* wage rates, reductions in fertility, and technological advances in household production are responsible for the huge increase in labor force participation of married women in this century. Changes in cultural and legal attitudes toward working women, as well as the social and economic disruptions brought about by two world wars and the Great Depression, also played a role.

---

[26] John Cogan, "Married Women's Labor Supply: A Comparison of Alternative Estimation Procedures," in Smith, editor, *Female Labor Supply: Theory and Estimation,* p. 113.

[27] U.S. Bureau of the Census, *Statistical Abstract of the United States,* Washington, DC: Government Printing Office, various issues.

[28] Joshua D. Angrist and William N. Evans, "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *American Economic Review* 88 (June 1998): 450–477.

A fascinating example is that unmarried young women living in states that granted them an early right to obtain oral contraceptives (i.e., the pill) without parental consent experienced a faster increase in labor force participation rates.[29] However, the evidence indicates that economic factors *do* matter and that a significant part of the increase in the labor force participation of married women can be understood in terms of the changing economic environment. It has been estimated that about 60 percent of the total growth in the female labor force between 1890 and 1980 can be attributed to the rising real wage of women.[30]

In recent years, technological changes in the labor market have allowed an increasing number of workers to do much of their work at home, further changing labor supply incentives. A recent study, in fact, reports that women who find it expensive to enter the labor market—such as women with small children—have strong incentives to use their home as their work base.[31] For example, only 15 percent of all women aged 25–55 who worked in a traditional "onsite" setting had children under the age of six. In contrast, 30 percent of "home-based" workers had children under the age of six. The prevalence of home-based work will likely rise as firms discover and adopt new technologies that allow them to outsource much of their work to other sites.

Many studies have attempted to estimate the responsiveness of women's hours of work to changes in the wage rate. Unlike the consensus estimate of the labor supply elasticity for prime-age men (that is, an elasticity on the order of -0.1), most studies of female labor supply find a *positive* relationship between a woman's hours of work and her wage rate, so substitution effects dominate income effects among working women. Recent studies that control for the selectivity bias arising from estimating labor supply models in the nonrandom sample of working women, however, tend to indicate that the size of the female labor supply elasticity may not be very large, perhaps on the order of 0.2.[32] A 10 percent increase in the woman's wage, therefore, increases her hours of work by about 2 percent.

Because of the huge changes in female labor supply witnessed in recent decades, there is a perception that female labor supply is more elastic than male labor supply. It is important to stress, however, that this perception is mostly due to the fact that female labor force

---

[29] Martha J. Bailey, "More Power to the Pill: The Effect of Contraceptive Freedom on Women's Life Cycle Labor Supply," *Quarterly Journal of Economics* 121 (February 2006): 289–320. See also Claudia Goldin, *Understanding the Gender Gap: An Economic History of American Women,* New York: Oxford University Press, 1990.

[30] Smith and Ward, "Time-Series Growth in the Female Labor Force"; Goldin, *Understanding the Gender Gap,* pp. 122–138; and Claudia Goldin, "The Role of World War II in the Rise of Women's Employment," *American Economic Review* 81 (September 1991): 741–756. More recent work finds the very rapid increase in labor force participation rates among successive cohorts of American women can only be explained if the rise in the real wage of women was accompanied by a reduction in the cost of raising children. See Orazio Attanasio, Hamish Low, and Virginia Sánchez-Marcosn, "Explaining Changes in Female Labor Supply in a Life-Cycle Model," *American Economic Review* 98 (September 2008): 1517–1552.

[31] Linda N. Edwards and Elizabeth Field-Hendrey, "Home-Based Work and Women's Labor Force Decisions," *Journal of Labor Economics* 20 (January 2002): 170–200.

[32] See Thomas Mroz, "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions," *Econometrica* 55 (July 1987): 765–800; Francine D. Blau and Lawrence M. Kahn, "Changes in the Labor Supply of Married Women: 1980–2000," *Journal of Labor Economics* 25 (July 2007): 393–438; and Bradley T. Haim, "The Incredible Shrinking Elasticities: Married Female Labor Supply, 1978–2002," *Journal of Human Resources* 42 (Fall 2007): 881–918.

participation rates are very responsive to changes in the wage. Among working women, however, there is growing evidence that women's hours of work, like those of men, are not very responsive to changes in the wage. Put differently, female labor supply mainly responds to economic factors at the margin of deciding whether or not to work, rather than at the margin of deciding how many hours to work once in the labor force.

The evidence also suggests that the labor force participation rates and hours of work of married women respond to changes in the husband's wage. A 10 percent increase in the husband's wage lowers the participation rate of women by 5.3 percentage points and reduces the hours that working wives allocate to the labor market by 1.7 percent. There is little evidence, however, that the husband's labor supply is affected by the wife's wage rate.[33] Overall, the empirical studies show some support for the notion that the family's labor supply decisions are jointly made by the various family members, with female labor supply being particularly responsive to changes in the husband's wage.

2-10   Policy Application: Welfare Programs and Work Incentives

The impact of income maintenance programs, such as Aid to Families with Dependent Children (AFDC) or Temporary Assistance for Needy Families (TANF), on the work incentives of recipients has been hotly debated since the days when the United States declared a war on poverty in the mid-1960s. In fact, much of the opposition to welfare programs was motivated by the conjecture that these programs encourage recipients to "live off the dole" and foster dependency on public assistance. The perception that welfare does not work and that the so-called War on Poverty was lost found a sympathetic ear among persons on all sides of the political spectrum and led to President Clinton's promise to "end welfare as we know it."[34] This political consensus culminated in the enactment of the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) in August 1996. The welfare reform legislation imposed lifetime limits on the receipt of various types of welfare programs, tightened eligibility requirements for most families, and mandated that many benefit-receiving families engage in work-related activities.

Cash Grants and Labor Supply

To illustrate how welfare programs can alter work incentives, let's begin by considering a simple program that grants eligible persons a cash grant. In particular, suppose that eligible persons (such as unmarried women with children) are given a cash grant of, say, $500 per month as long as they remain outside the labor force. If these persons enter the labor market, the government officials immediately assume that the women no longer need public assistance and the women are dropped from the welfare rolls (regardless of how much they earned).

---

[33] Orley Ashenfelter and James J. Heckman, "The Estimation of Income and Substitution Effects in a Model of Family Labor Supply," *Econometrica* 42 (January 1974): 73–85; and Shelly Lundberg, "Labor Supply of Husbands and Wives: A Simultaneous Equation Approach," *Review of Economics and Statistics* 70 (May 1988): 224–235.

[34] Charles Murray, *Losing Ground: American Social Policy, 1950–1980,* New York: Basic Books, 1984; and David T. Ellwood, *Poor Support: Poverty in the American Family,* New York: Basic Books, 1988. A survey of the academic studies that assess the impact of these programs is given by Robert Moffitt, "Incentive Effects of the U.S. Welfare System: A Review," *Journal of Economic Literature* 30 (March 1992): 1–61.

FIGURE 2-14   **Effect of a Cash Grant on Work Incentives**

A take-it-or-leave-it cash grant of $500 per month moves the worker from point *P* to point *G,* and encourages the worker to leave the labor force.
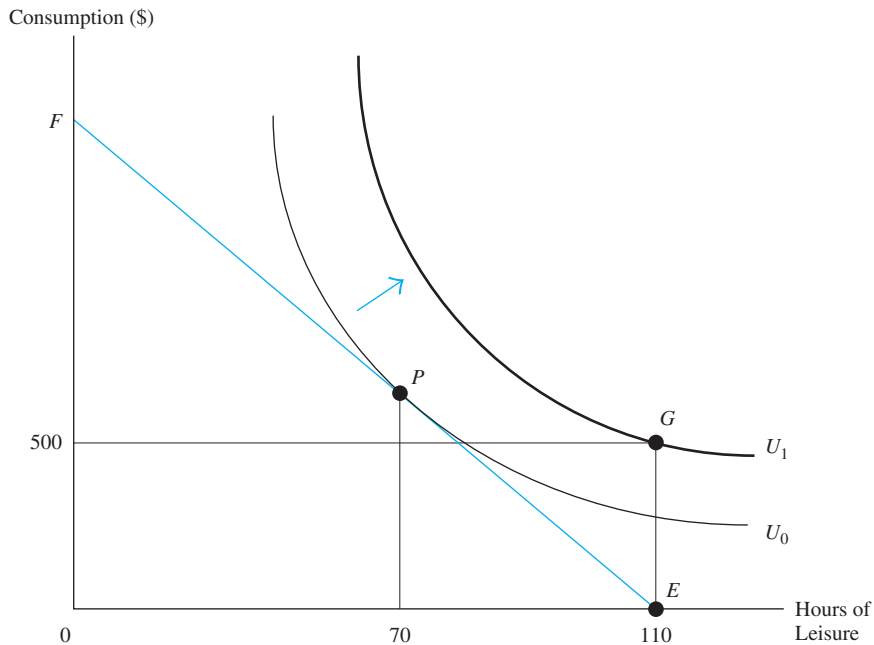
The impact of the cash grant on work incentives is illustrated in Figure 2-14. In the absence of the program, the budget line is given by *FE* and leads to an interior solution at point *P,* in which the person consumes 70 hours of leisure and works 40 hours.

For simplicity, assume that the woman does not have any nonlabor income. The introduction of a cash grant of $500 to nonworkers then introduces point *G* into the opportunity set. At this point, the woman can purchase $500 worth of consumption goods if she participates in the welfare program and does not work. Once the woman enters the labor market, however, the welfare grant is taken away and the opportunity set switches back to the original budget line *FE.*

The existence of the cash grant at point *G* can greatly reduce work incentives. As drawn, the woman attains a higher level of utility by choosing the corner solution at point *G* (that is, the welfare solution) than by choosing the interior solution at point *P* (that is, the work solution).

This type of "take-it-or-leave-it" cash grant can induce many workers to drop out of the labor market. In fact, it should be clear that low-wage women are most likely to choose the welfare solution. An improvement in the endowment point (from point *E* to point *G*) increases the worker's reservation wage, reducing the likelihood that a low-wage person will enter the labor market.

It is important to emphasize that welfare programs do not lower the labor force participation rates of low-wage workers because these workers lack a "work ethic." After all, we have implicitly assumed that the preferences of low-wage workers (as represented by the family of indifference curves) are identical to the preferences of high-wage workers.

Rather, the welfare program reduces the work incentives of low-wage workers because it is these workers who are most likely to find that the economic opportunities provided by the welfare system are better than those available in the labor market.

The Impact of Welfare on Labor Supply

In view of the extreme disincentive effects of the program illustrated in Figure 2-14, social assistance programs typically allow welfare recipients to remain in the labor force. Although welfare recipients can work, the amount of the cash grant is often reduced by some specified amount for every dollar earned in the labor market. Prior to 1996, for example, the AFDC grant was reduced by 67 cents for every dollar that the woman earned in the labor market (during the first four months that the woman was on welfare).[35]

It is instructive to describe with a numerical example how this type of welfare program alters the person's opportunity set. Suppose that, if the woman does not work at all and goes on welfare, her monthly income is $500 (assuming that she does not have any other non-labor income). For the purposes of this example, suppose that the government takes away 50 cents from the cash grant for every dollar earned in the labor market. This means that, if the woman works one hour at a wage of $10, her labor earnings increase by $10 but her grant is reduced by $5. Her total income, therefore, is $505. If she decides to work two hours, her labor earnings are $20 but her grant is reduced by $10. Total income would then be $510. Every additional hour of work increases income by only $5. Under the guise of reducing the size of the welfare grant, the government is actually taxing the welfare recipient's wage at a 50 percent rate. Therefore, it becomes important to differentiate between the woman's *actual* wage rate (which is $10 an hour) and the *net* wage (which is only $5 an hour).

Figure 2-15 illustrates the budget line created by this type of welfare program. In the absence of the program, the budget line is given by *FE* and the woman would choose the consumption bundle given by point *P.* She would then consume 70 hours of leisure and work 40 hours.

The welfare program shifts the budget line in two important ways. Because of the $500 monthly grant when the woman does not work, the endowment point changes from point *E* to point *G.* The program also changes the slope of the budget line. We have seen that the reduction of the grant by 50 cents for every dollar earned in the labor market is equivalent to a 50 percent tax on her earnings. The relevant slope of the budget line, therefore, is the net wage rate. Hence the welfare program cuts the (absolute value of the) slope by half, from $10 to $5. The budget line associated with the welfare program is then given by *HG.*

As drawn, when given the choice between the budget line *FE* and the budget line *HG,* the woman opts for the welfare system and chooses the consumption bundle given by point *R.* She consumes 100 hours of leisure and works 10 hours. Even this liberal "workfare" program, therefore, seems to have work disincentives because she works fewer hours than she would have worked in the absence of welfare.

---

[35] The taxation scheme implicit in the pre-1996 AFDC program was actually quite peculiar. During the first four months of a welfare spell, the welfare recipient was allowed to keep the first $90 earned per month (this amount was called the "earnings disregard"), but any additional earnings were taxed at a 67 percent tax rate. After being on welfare for four months, the earnings disregard was still $90 per month, but additional earnings were taxed at a 100 percent rate. An exhaustive description of the parameters of all means-tested entitlement programs in the United States is given by the Committee on Ways and Means, U.S. House of Representatives, *Overview of Entitlement Programs, Green Book,* Washington, DC: Government Printing Office, various issues.

FIGURE 2-15   **Effect of a Welfare Program on Hours of Work**

A welfare program that gives the worker a cash grant of $500 and imposes a 50 percent tax on labor earnings reduces work incentives. 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.
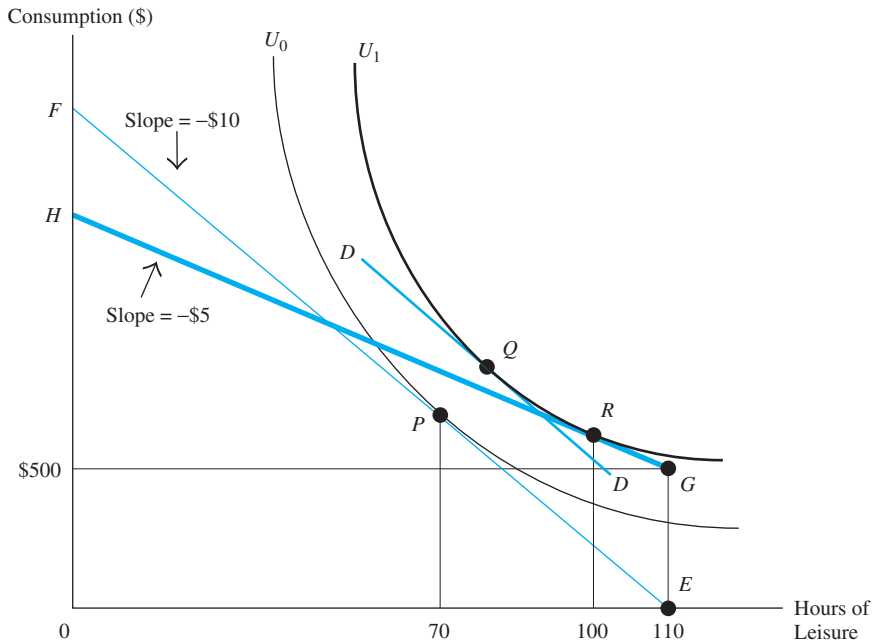
In fact, we can demonstrate that a welfare program that includes a cash grant and a tax on labor earnings *must* reduce hours of work. In particular, point *R* must be to the right of point *P*. To see why, draw a hypothetical budget line parallel to the pre-welfare budget line, but tangent to the new indifference curve. This line is labeled *DD* in Figure 2-15. It is easy to see that the move from point *P* to point *Q* is an income effect and represents the impact of the cash grant on hours of work. This income effect increases the demand for leisure. In other words, point *Q* must be to the right of point *P*.

The move from point *Q* to point *R* represents the substitution effect induced by the 50 percent tax on labor earnings, and point *R* must be to the right of point *Q*. The tax cuts the price of leisure by half for welfare recipients. As a consequence, the welfare recipient will demand even more leisure.

This stylized example vividly describes the work incentive problems introduced by welfare programs. If our model adequately represents how persons make their work decisions, it is impossible to formulate a relatively generous welfare program without substantially reducing work incentives. Awarding cash grants to recipients, as welfare programs unavoidably do, reduces both the probability of a person working and the number of hours worked by those who remain on the job. In addition, efforts to recover some of the grant money from working welfare recipients effectively impose a tax on work activities. This tax reduces the price of leisure and further lowers the number of hours that the welfare recipient will work.

The study of how welfare programs affect work incentives shows how the basic framework provided by the neoclassical model of labor-leisure choice is a point of departure that can be used to analyze more complex situations. By specifying in more detail how a person's opportunities are affected by government policies, we can easily adapt the model to analyze important social questions. The beauty of the economic approach is that we do not need different models to analyze labor supply decisions under alternative government policies or social institutions. In the end, we are always analyzing the *same* model—how workers allocate their limited time and money so as to maximize their utility—but we keep feeding the model more detail about the person's opportunity set.

Welfare Reform and Labor Supply

As we saw earlier, the theory predicts that welfare programs create work disincentives. In fact, many of the studies that studied the impact of the pre-1996 welfare programs typically found that the AFDC program reduced labor supply by 10 to 50 percent from the level of work effort that would have been found in the absence of the program, and the values of the labor supply elasticities generally fell in line with the consensus estimates described above.[36]

On August 22, 1996, President Clinton signed into law the welfare reform legislation that fundamentally changed the welfare system of the United States. A key provision in the legislation gave states a great deal of freedom in setting eligibility rules and benefit levels for many assistance programs.[37] For example, California now allows a TANF recipient to earn up to $225 per month without affecting the size of the welfare benefit, but any additional earnings are taxed at a 50 percent rate. In contrast, Illinois taxes all labor earnings at a 33 percent rate, while Mississippi applies a 100 percent tax rate on any labor earnings above $90 per month.

Many studies have used this variation across states to determine the impact of welfare programs on labor supply and many other variables, including the size of the welfare population itself. One difficult problem with the studies that evaluate the welfare reform legislation is that the period immediately following the enactment of PRWORA coincided with a historic economic boom in the United States. As a result, it has been difficult to determine how much of the decline in the size of the welfare caseload (from 4.4 million families receiving TANF in August 1996 to 2.2 million in June 2000) can be attributed to the economic boom and how much can be attributed to the changes in welfare policy.[38]

---

[36] For example, see Hilary Williamson Hoynes, "Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation under AFDC-UP," *Econometrica* 64 (March 1996): 295–332. This literature is reviewed by Alan B. Krueger and Bruce D. Meyer, "Labor Supply Effects of Social Insurance," in Alan Auerbach and Martin Feldstein, editors, *Handbook of Public Economics,* Vol. 4, Amsterdam: North-Holland, 2002; and Robert A. Moffitt, "Welfare Programs and Labor Supply," in Alan Auerbach and Martin Feldstein, editors, *Handbook of Public Economics,* Vol. 4, Amsterdam: North-Holland, 2002.

[37] For a detailed discussion of the state differences in the TANF program and of the available research, see Robert A. Moffitt, "The Temporary Assistance for Needy Families Program," in Robert A. Moffitt, *Means-Tested Transfer Programs in the United States,* Chicago: University of Chicago Press, 1993, pp. 291–363.

[38] Robert F. Schoeni and Rebecca Blank, "What Has Welfare Reform Accomplished? Impacts on Welfare Participation, Employment, Poverty, Income, and Family Structure," National Bureau of Economic Research Working Paper No. 7627, March 2000; Jeffrey Grogger, "The Effects of Time Limits, the EITC, and Other Policy Changes on Welfare Use, Work, and Income among Female-Headed Families," *Review of Economics and Statistics* 85 (May 2003): 394–408.

Many states have conducted large-scale experiments. In the typical experiment, a group of randomly chosen families is offered a particular set of program parameters and benefits, while other families are offered a different set. By investigating the variation in labor supply among the different groups of families, it is possible to determine if labor supply responds to the financial incentives implied by the program parameters. These experiments often confirm the theoretical predictions.[39] One well-known experiment, the Minnesota Family Investment Program, allowed women to keep some of the cash benefits even if their earnings were relatively high (about 140 percent of the poverty line). The results of this experiment indicated that reducing the tax on labor earnings indeed encouraged the welfare recipients to work more.

There also has been a lot of interest in determining the impact of "time limits" on welfare participation. A key provision of PRWORA limits the amount of time that families can receive federal assistance to 60 months over their lifetimes, and many states have used their authority to set even shorter time limits.

The presence of time limits introduces interesting strategic choices for an eligible family: a family may choose to "bank" its benefits in order to maintain eligibility further into the future. Federal law permits welfare payments only to families that have children younger than 18 years of age. As a result, the family's choice of whether to receive assistance today (and use up some of its 60 eligible months) or to save its eligibility for a later period depends crucially on the age of the youngest child. Families with older children might as well use up their benefits now since it is unlikely that they can qualify for benefits some years into the future. In contrast, families with younger children have a longer time span over which they must allow for the possibility that they will require assistance, and they have an incentive not to use up the 60 months of lifetime benefits too soon.

The evidence strongly confirms this interesting insight. Time limits have the greatest effect on welfare participation rates of families with small children. All other things equal, the presence of time limits reduces the welfare participation of families where the youngest child is 3 years old by about 8 percentage points relative to the welfare participation of families where the youngest child is 10 years old.[40]

2-11   Policy Application: The Earned Income Tax Credit

An alternative approach to improving the economic status of low-income persons is given by the Earned Income Tax Credit (EITC). This program began in 1975 and has been expanded substantially since. By 2007, the EITC was the largest cash-benefit entitlement program in the United States, granting nearly $40 billion to low-income households.

To illustrate how the EITC works, consider a household composed of a working mother with two qualifying children. In 2005, for example, this woman could claim a tax credit of up to 40 percent of her earnings as long as she earned less than $11,000 per year, resulting in

---

[39] See Jeffrey Grogger; Lynn A. Karoly, and Jacob Alex Klerman, *Consequences of Welfare Reform: A Research Synthesis,* Santa Monica, CA: The Rand Corporation, July 2002; and Rebecca Blank, "Evaluating Welfare Reform in the U.S.," *Journal of Economic Literature* 40 (December 2002): 1105–1166.

[40] Jeffrey Grogger, "Time Limits and Welfare Use," *Journal of Human Resources* 39 (Spring 2004): 405–424; and Jeffrey Grogger and Charles Michalopoulos, "Welfare Dynamics under Time Limits," *Journal of Political Economy* 111 (June 2003): 530–554.

FIGURE 2-16 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 $11,000. The credit is capped at $4,400. The worker receives this maximum amount as long as she earns between $11,000 and $14,370. 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 $14,370 and $35,263.
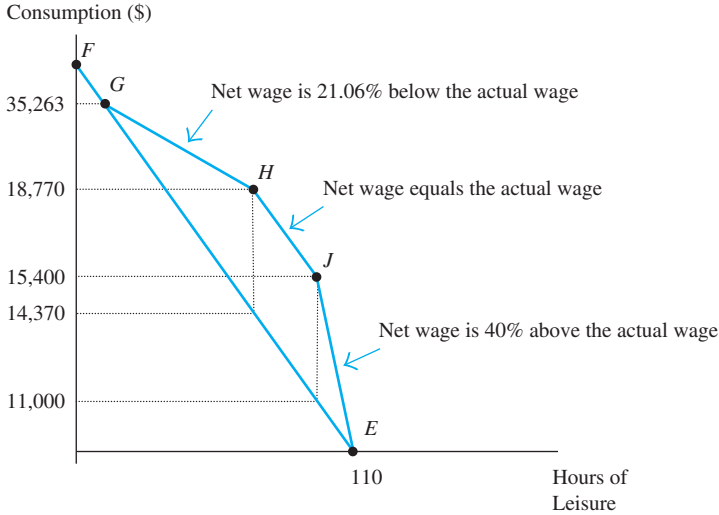
a maximum credit of $4,400. This maximum credit would be available as long as she earned between $11,000 and $14,370. After reaching the $14,370 threshold, the credit would begin to be phased out. In particular, each additional dollar earned reduces the credit by 21.06 cents. This formula implies that the credit completely disappears once the woman earns $35,263.

Figure 2-16 illustrates how the EITC introduces a number of "kinks" into the worker's opportunity set. The figure assumes that the worker does not have any nonlabor income. In the absence of the EITC, the worker faces the straight budget line given by *FE*. The EITC changes the net wage associated with an additional hour of work. As long as the worker earns less than $11,000 per year, the worker can claim a tax credit of up to 40 percent of earnings. Suppose, for instance, that the wage rate is $10 an hour and that the worker decides to work only one hour during the entire year. She can then file a tax return that would grant her a $4 tax credit. Therefore, the EITC implies that the worker's net wage is $14, a 40 percent raise. This 40 percent tax credit makes the budget line steeper, as illustrated by the segment *JE* in Figure 2-16.

If the woman earns $11,000, she receives the maximum tax credit, or $4,400. In fact, she is eligible for this maximum credit as long as she earns anywhere between $11,000 and $14,370. As long as the worker is in this range, therefore, the EITC does not change the net wage. It simply generates an increase in the worker's income of $4,400—as illustrated by the segment *HJ* in Figure 2-16, which illustrates that the EITC generates a pure income effect in this range of the program.

Once the worker's annual earnings exceed $14,370, the EITC is phased out at a rate of 21.06 cents for every dollar earned. Suppose, for example, that the worker earns exactly $14,370 and decides to work an additional hour at $10 an hour. The tax credit is then cut

back by about $2.11, implying that the worker's net wage is only $7.89 an hour. The EITC, therefore, acts like a wage cut, flattening out the budget line, as illustrated by segment *GH* in Figure 2-16. Once the worker earns $35,263 during the year, she no longer qualifies for the EITC and her budget line reverts back to the original budget line (as in segment *FG*).

This detailed illustration of how the EITC works illustrates how government programs change the worker's opportunity set, creating strangely shaped budget lines with a number of kinks. These kinks can have important effects on the worker's labor supply decision.

So how does the EITC affect labor supply? The various panels of Figure 2-17 illustrate a number of possibilities. In Figure 2-17*a,* the worker would not be in the labor force in the absence of the EITC program (she maximizes her utility by being at the endowment point *P*). The increase in the net wage associated with the EITC draws the woman into the labor force, and she maximizes her utility by moving to point *R*. The reason for the increased propensity to work should be clear from our previous discussion. The EITC increases the net wage for nonworkers, making it more likely that the labor market can match their reservation wages and, hence, encouraging these persons to join the labor force. The theory, therefore, has a clear and important prediction: the EITC should increase the labor force participation rate in the targeted groups.

In Figure 2-17*b,* the person would be in the labor force even if the EITC were not in effect (at point *P*). This worker's annual income implies that the EITC generates an income effect—without affecting the net wage. The worker maximizes her utility by moving to point *R,* and she would be working fewer hours.

Finally, in Figure 2-17*c,* the person would work a large number of hours in the absence of the EITC (at point *P*). The EITC cuts her net wage, and she maximizes her utility by cutting hours and moving to the kink at point *R*.

The theory, therefore, suggests that the EITC has two distinct effects on labor supply. First, the EITC increases the number of labor force participants. Because the tax credit is granted only to persons who work, more persons will enter the labor force to take advantage of this program. Second, the EITC may change the number of hours worked by persons who would have been in the labor force even in the absence of the program. As drawn in the various panels of Figure 2-17, the EITC motivated workers to work fewer hours—but the change in the net wage generates both income and substitution effects and the impact of the EITC on hours worked will depend on the relative importance of these two effects.

The available evidence confirms the theoretical prediction that the EITC draws many new persons into the labor force.[41] Some of this evidence is summarized in Table 2-5. The Tax Reform Act of 1986 substantially expanded the benefits available through the EITC. The theory suggests that this legislative change should have increased the labor force participation rates of the targeted groups. Consider the population of unmarried women in the United States. Those who have at least one child potentially qualify for the EITC (depending on how much they earn), whereas those without children do not qualify. Table 2-5 shows that the labor force participation rate of the eligible women increased from 72.9 percent to 75.3 percent before and after the 1986 tax reform went into effect, an increase of 2.4 percentage points.

---

[41] V. Joseph Hotz and John Karl Scholz, "The Earned Income Tax Credit," in Robert A. Moffitt, editor, *Means-Tested Transfer Programs in the United States,* Chicago: University of Chicago Press, 2003; and Nada Eissa and Hilary W. Hoynes, "Behavioral Responses to Taxes: Lessons from the EITC and Labor Supply," *Tax Policy and the Economy* 20(2006): 74–110.

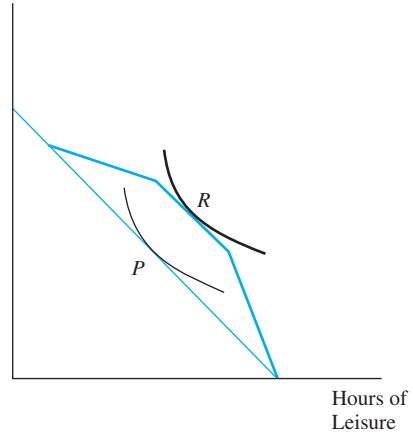## FIGURE 2-17 The Impact of the EITC on Labor Supply

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

Consumption ($)   Consumption ($)

P   Hours of   Hours of
            Leisure   Leisure
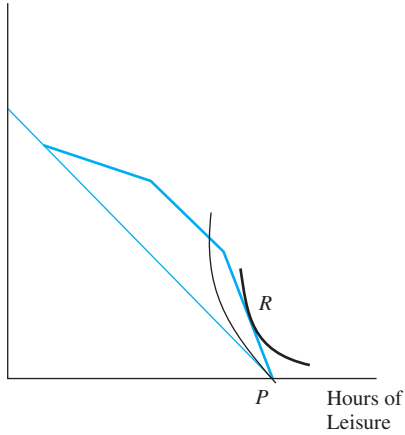
(*a*) EITC Draws Worker into Labor Market   (*b*) EITC Reduces Hours of Work

Consumption ($)
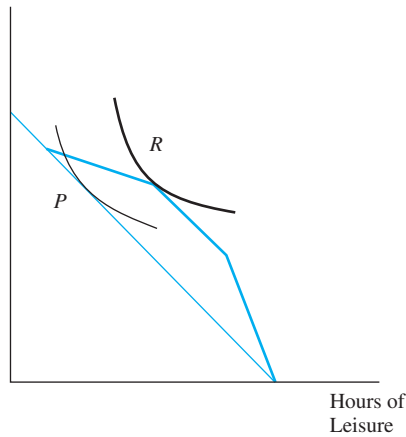
Hours of
Leisure

(*c*) EITC Reduces Hours of Work

TABLE 2-5   **The Impact of the Earned Income Tax Credit on Labor Force Participation**

Source: Nada Eissa and Jeffrey B. Liebman, "Labor Supply Response to the Earned Income Tax Credit," *Quarterly Journal of Economics* 111 (May 1996): 617.

	Participation Rate before Legislation (%)	Participation Rate after Legislation (%)	Difference (%)	Difference-in-Differences (%)
Treatment group—eligible for the EITC:				
Unmarried women with children	72.9	75.3	2.4	
Control group—not eligible for the EITC:				2.4
Unmarried women without children	95.2	95.2	0.0	

Before one can conclude that this change in labor force participation rates can be attributed to the EITC, one must consider the possibility that other factors might account for the 2.4 percentage point increase in labor force participation rates observed during that period. A booming economy, for instance, could have easily drawn more women into the labor market even in the absence of the EITC. Or there could exist long-run demographic and social trends that might account for the increasing propensity for these women to enter the labor force.

As in the typical experiment conducted in the natural sciences, we need a "control group"—a group of workers that would have experienced the same types of macroeconomic or demographic changes but that were not "injected" with the benefits provided by the EITC. Such a group could be the group of unmarried women without children. It turns out that their labor force participation did not change at all as a result of the Tax Reform Act of 1986—it stood at 95.2 percent both before and after the tax reform legislation.

The impact of the EITC on labor force participation, therefore, can be calculated by comparing the trend in the "treatment group"—the unmarried women with children—with the trend in the "control group"—the unmarried women without children. The labor force participation rate changed by 2.4 percentage points in the treatment group and by 0 percentage points in the control group. One can then estimate the net impact of the EITC on labor force participation by taking a "difference-in-differences": 2.4 percentage points minus 0 percentage points, or 2.4 percentage points.

This methodology for uncovering the impact of specific policy changes or economic shocks on labor market outcomes is known as the **difference-in-differences estimator** and has become very popular in recent years. The approach provides a simple way of measuring how particular events can alter labor market opportunities. At the same time, however, it is important to recognize that the validity of the conclusion depends crucially on our having chosen a correct control group that nets out the impact of *all* other factors on the trends that we are interested in.[42]

It is worth concluding by remarking briefly on the labor supply consequences of the two distinct approaches that we have discussed for subsidizing disadvantaged workers. The typical welfare program uses a "cash grant"—granting income grants to persons who do not or cannot work. As we have seen, these grants can greatly reduce work incentives

[42] Marianne Bertrand, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119 (February 2004): 249–275.

and make it more likely that program participants do not enter the labor force. The earned income tax credit, in contrast, subsidizes work. It does not provide a cash grant, and instead increases the net wage for nonworkers who enter the labor force. As a result, it can greatly increase work incentives and make it more likely that eligible recipients work.

2-12   Labor Supply over the Life Cycle

Up to this point, our model of labor supply analyzes the decisions of whether to work and how many hours to work from the point of view of a worker who allocates his time in a single time period and who ignores the fact that he will have to make similar choices continuously over many years. In fact, because consumption and leisure decisions are made over the entire working life, workers can "trade" some leisure time today in return for additional consumption tomorrow. For instance, a person who devotes a great deal of time to his job today can save some of the additional earnings and use these savings to increase his consumption of goods in the future.

As we will see in Chapter 6, a great deal of evidence suggests that the typical worker's age-earnings profile—the worker's wages over the life cycle—has a predictable path: wages tend to be low when the worker is young; they rise as the worker ages, peaking at about age 50; and the wage rate tends to remain stable or decline slightly after age 50. The path of this typical age-earnings profile is illustrated in Figure 2-18*a.* This age-earnings profile implies that the price of leisure is relatively low for younger and older workers and is highest for workers in their prime-age working years.

Consider how the worker's labor supply should respond to the wage increase that occurs between ages 20 and 30, or to the wage decline that might occur as the worker nears retirement age. It is important to note that these types of wage changes are part of the aging process *for a given worker.* A change in the wage along the worker's wage profile is called an "evolutionary" wage change, for it indicates how the wages of a particular worker evolve over time. It is crucial to note that an evolutionary wage change has no impact whatsoever on the worker's total *lifetime income.* The worker fully expects his wage to go up as he matures and to go down as he gets closer to retirement. As a result, an evolutionary wage change alters the price of leisure—but does not alter the value of the total opportunity set available to the worker over his life cycle. To be more precise, suppose we know that our life cycle age-earnings profile takes on the precise shape illustrated in Figure 2-18*a.* The fact that our wage rises slightly from age 37 to age 38 or declines slightly from age 57 to age 58 does not increase or decrease our lifetime wealth. We already expected these evolutionary wage changes to occur and they have already been incorporated in the calculation of lifetime wealth.

Suppose then that the wage falls as a worker nears retirement age, and consider the following question: Would the worker be better off by working a lot of hours at age 50 and consuming leisure in his sixties, or would the worker be better off by working relatively few hours at age 50 and devoting a great deal of time to his job in his sixties?

The worker will clearly find it worthwhile to work more hours at age 50, invest the money, and buy consumption goods and leisure at some point in the future when the wage is lower and leisure is not as expensive. After all, this type of labor supply decision would increase the worker's lifetime wealth; it gives him a much larger opportunity set than would be available if he were to work many hours in his sixties (when the wage is low) and consume many hours of leisure in his fifties (when the wage is high).

FIGURE 2-18 **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 over the life cycle 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.
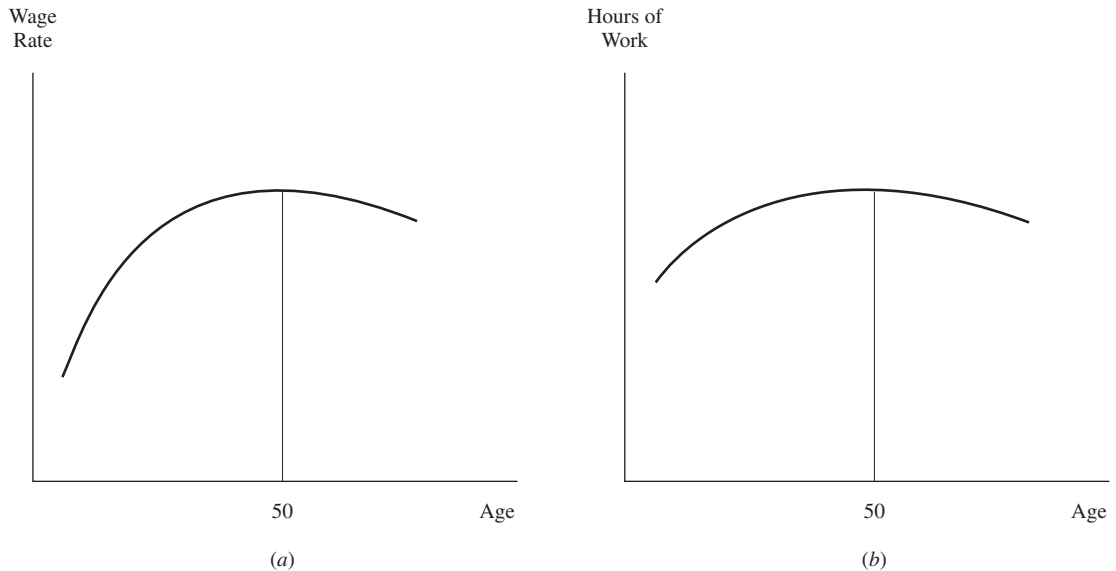

A very young worker faces the same type of situation. His wage is relatively low—and he will find it optimal to consume leisure activities when he is very young, rather than in his thirties and forties, when the price of those leisure activities will be very high. The argument, therefore, suggests that we will generally find it optimal to concentrate on work activities in those years when the wage is high and to concentrate on leisure activities in those years when the wage is low.[43]

In the end, this approach to life cycle labor supply decisions implies that hours of work and the wage rate should move together over time *for a particular worker,* as illustrated in Figure 2-18*b*. This implication differs strikingly from our earlier conclusion that a wage increase generates both income and substitution effects, and that there could be a negative relationship between wages and hours of work if income effects dominate. This important difference between the models (that is, the one-period "static" model considered in the previous sections and the life cycle model presented here) arises because the two models mean very different things by a change in the wage. In the one-period model, an increase in the wage expands the worker's opportunity set and hence creates an income effect that increases the demand for leisure. In the life cycle model, an evolutionary wage change—the wage change that workers expect as they age—does not change the total lifetime income available to a *particular* worker, and leaves the lifetime opportunity set intact.

[43] A detailed exposition of the model is given by James J. Heckman, "Life Cycle Consumption and Labor Supply: An Explanation of the Relationship between Income and Consumption over the Life Cycle," *American Economic Review* 64 (March 1974): 188–194.

FIGURE 2-19 **Hours of Work over the Life Cycle for Two Workers with Different Wage Paths**

Joe's wage exceeds Jack's at every age. Although both Joe and Jack work more hours when the wage is high, Joe works more hours than Jack only if the substitution effect dominates. If the income effect dominates, Joe works fewer hours than Jack.


(a)

(b)
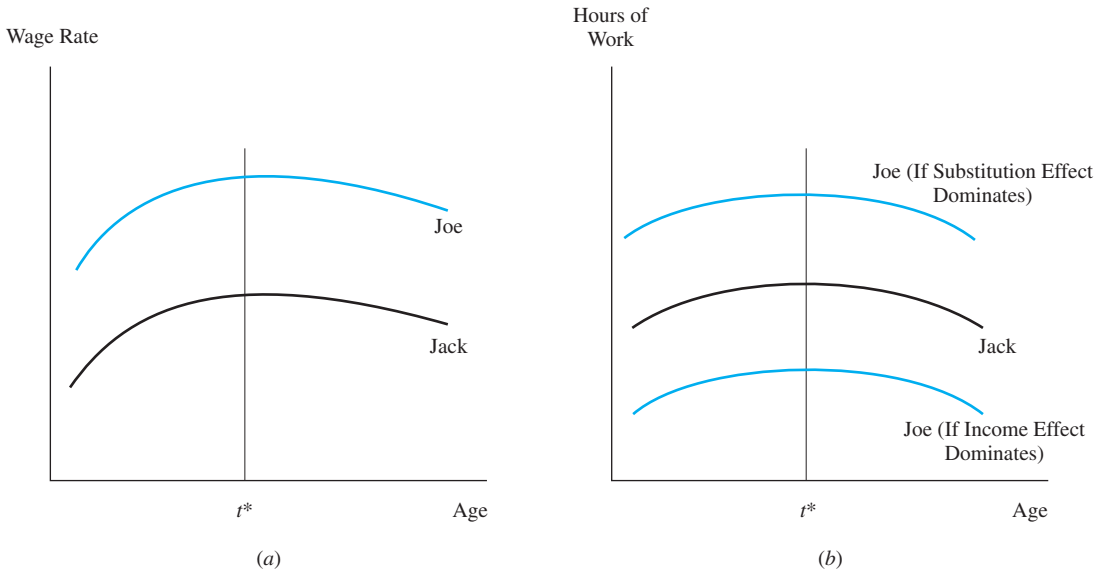
In contrast, if we were to compare two workers, say Joe and Jack, with different age-earnings profiles, the difference in hours of work between these two workers would be affected by both income and substitution effects. As illustrated in Figure 2-19a, Joe's wage exceeds Jack's at every age. Both Joe and Jack should work more hours when wages are high. Their life cycle profiles of hours of work are illustrated in Figure 2-19b. We do not know, however, which of the two workers allocates more hours to the labor market. In particular, even though Joe has a higher wage and finds leisure to be a very expensive commodity, he also has a higher lifetime income and will want to consume more leisure. The difference in the level of the two wage profiles, therefore, generates an income effect. If these income effects are sufficiently strong, Joe's hours-of-work profile will lie below Jack's; if substitution effects dominate, Joe will work more hours than Jack at every age.

The life cycle approach suggests a link not only between wages and hours of work, but also between wages and labor force participation rates. As we saw earlier in the chapter, the labor force participation decision depends on a comparison of the reservation wage to the market wage. In each year of the life cycle, therefore, the worker will compare the reservation wage to the market wage. Suppose initially that the reservation wage is roughly constant over time. The person is then more likely to enter the labor market in periods when the wage is high. As a result, participation rates are likely to be low for young workers, high for workers in their prime working years, and low again for older workers.

The participation decision, however, also depends on how reservation wages vary over the life cycle. The reservation wage measures the bribe required to enter the labor market. For instance, the presence of small children in the household increases the value of time in
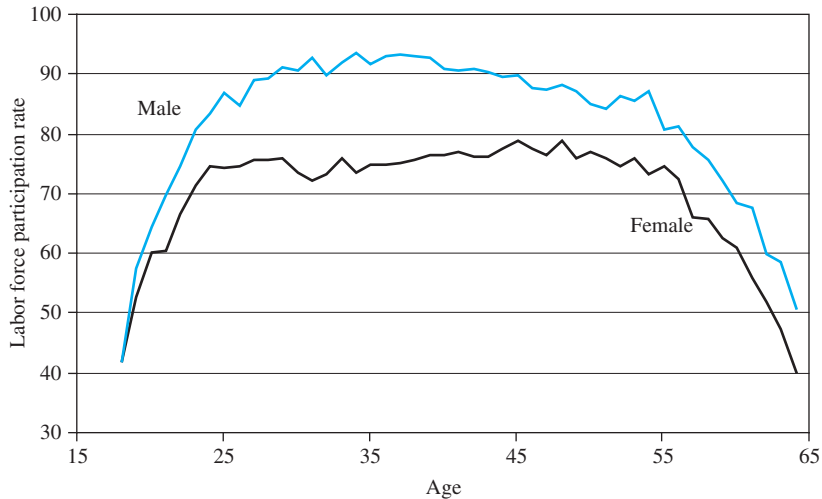
the nonmarket sector for the person most responsible for child care and, hence, also would increase the reservation wage. Therefore, it is not surprising to find that some married women participate in the labor force intermittently. They work prior to the arrival of the first child, withdraw from the labor market when the children are small and need full-time care, and return to the labor market once the children enroll in school.

The key implication of the analysis can be easily summarized: A person will work few hours in those periods of the life cycle when the wage is low and will work many hours in those periods when the wage is high. The evidence on age-earnings profiles suggests that the wage is relatively low for young workers, increases as the worker matures and accumulates various types of skills, and then may decline slightly for older workers. The model then suggests that the profile of hours of work over the life cycle will have exactly the same shape as the age-earnings profile: hours of work increase as the wage rises and decline as the wage falls. The theoretical prediction that people allocate their time over the life cycle so as to take advantage of changes in the price of leisure is called the **intertemporal substitution hypothesis.**

Evidence

The available evidence suggests that both labor force participation rates and hours of work respond to evolutionary wage changes. Figure 2-20 illustrates the relationship between labor force participation rates and age in the United States. Male participation rates peak when men are between 25 and 45 years old and begin to decline noticeably after age 45. In contrast, female participation rates, probably because of the impact of child-raising activities on the participation decision, do not peak until women are around 45 years old.

Overall, the trends illustrated in the figure are consistent with the theoretical prediction that participation rates should be highest when the wage is high (that is, when workers are in their thirties and forties). The decline in labor force participation rates
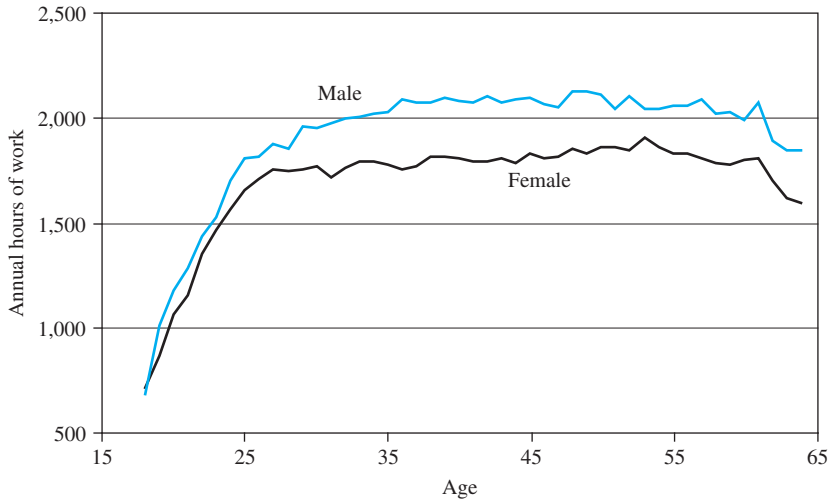
observed after age 55, however, is much too steep to be explained by the wage decline that is typically observed as workers near retirement age. The rapid decline in participation rates at older ages may be health related and, as we will see later in this chapter, also may be attributable to the work disincentive effects of various retirement and disability insurance programs.

Figure 2-21 illustrates the actual relationship between hours of work and age. As with participation rates, hours of work among working men rise rapidly until about age 30, peak at ages 35 to 45, and begin to decline at age 50. During the prime working years, men work about 2,100 hours annually. In contrast, hours of work among working women do not peak until age 50 (probably because some younger women work in part-time jobs while they have small children in the household).

Many studies have attempted to estimate the responsiveness of hours of work to changes in the wage over the life cycle.[44] These studies typically use a longitudinal sample of workers (that is, a data set where each person in the sample is followed over time) to estimate how a given worker adjusts his or her hours of work to the evolutionary wage changes that occur as the worker ages. The intertemporal substitution hypothesis implies that the correlation between changes in hours of work and changes in the wage should be positive: As a worker ages, an increase in the wage rate should increase hours of work.

---

[44] Thomas E. MaCurdy, "An Empirical Model of Labor Supply in a Life-Cycle Setting," *Journal of Political Economy* 89 (December 1981): 1059–1085. See also Joseph G. Altonji, "Intertemporal Substitution in Labor Supply: Evidence from Micro Data," *Journal of Political Economy* 94 (June 1986, Part 2): S176–S215; Joseph V. Hotz, Kinn Kydland, and Guilherme Sedlacek, "Intertemporal Preferences and Labor Supply," *Econometrica* 56 (March 1988): 335–360; and Casey Mulligan, "Substitution over Time: Another Look at Life Cycle Labor Supply," *NBER Macroeconomics Annual* 13 (1998): 75–134.

Taxi drivers in New York City typically pay a fixed fee to lease their cab for a prespecified period, such as a day or week. The driver is responsible for buying gas and for some of the car maintenance. As part of the leasing contract, the cabbie can keep whatever fare income he generates as he cruises the city streets. Every time he leases a cab, therefore, he faces an important labor supply decision: How long should he keep on looking for additional fares?

The work shift of a typical Manhattan cabbie surveyed in a recent study lasted 6.9 hours, of which only about 4.6 hours were actually spent driving a passenger. The rest of the time was spent cruising for a fare or taking a break. The total income during the shift was $161, so that the average hourly wage rate was around $23.

This average wage rate, however, probably masks a great deal of variation in the rewards to working an additional hour. The marginal wage rate probably depends greatly on the weather and on the time of the day and day of the week. For example, there may be many potential

passengers on a rainy Friday afternoon, as New Yorkers leave their offices early to prepare for the weekend.

The theory of intertemporal labor substitution implies that the typical cabbie should be willing to work a longer shift when he expects the city streets to be busy and full of potential passengers and to take leisure on those hours and days that are expected to be slower. It is not surprising, therefore, that there are relatively few cabs cruising the streets at 2 a.m. on a Monday morning. In fact, a recent study shows that cabbies respond to the changed economic situations during the day and during the week in a way that is consistent with the theory: They drive a longer shift when the marginal wage rate is higher.

Sources: Henry S. Farber, "Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers," *Journal of Political Economy* 113 (February 2005): 46–82; and Colin Camerer, Linda Babcock, George Loewenstein, and Richard Thaler, "Labor Supply of New York City Cabdrivers: One Day at a Time," *Quarterly Journal of Economics* 112 (May 1997): 407–441.

The data illustrated in Figure 2-21 clearly indicate that hours of work increase early on in the life cycle and decline as retirement age approaches. The data, however, also reveal that hours of work are "sticky" over a long stretch of the working life. For example, annual hours worked by men barely budge between the ages of 35 and 50, despite the fact that the wage rises substantially during this period. Because hours of work tend to be sticky, many studies conclude that the response of hours of work to evolutionary wage changes is small: a 10 percent increase in the wage leads to less than a 1 percent increase in hours of work. Therefore, labor supply over the life cycle (as defined by hours of work per year) may not be very responsive to changes in the wage.[45]

Estimation of Life Cycle Models

The estimation of the intertemporal labor supply elasticity—the crucial parameter that determines how hours of work evolve in the life cycle model of the labor-leisure choice—helped introduce what has become a very useful econometric technique into the labor

[45] It is important to stress, however, that there is a lot of debate over the validity of this conclusion. The magnitude of the labor supply response to life cycle changes in the wage (called the *intertemporal elasticity of substitution*) has important implications in macroeconomics. Some macroeconomic models require sizable intertemporal elasticities to explain the behavior of employment over the business cycle. As a result, there is heated disagreement over the evidence suggesting that the intertemporal elasticity is small.

economics literature.[46] The economic model states that we should be tracking a specific individual over the lifetime so that we can observe how his hours of work change from year to year as a response to year-to-year wage changes. Suppose that we have a longitudinal data set that allows us to observe a particular worker i twice, say, at the ages of 40 and 41. Let H_{it} give his hours of work at age t, and w_{it} gives his wage rate at that age. It is easy to see that one can difference the data for each individual estimate the following regression model across the sample of different workers:

$$\Delta H_{it} = \sigma \Delta w_{it} + \text{Other variables} \qquad \textbf{(2-14)}$$

where ΔH_{it} gives the year-to-year change in hours of work and Δw_t gives the year-to-year change in the worker's wages. The coefficient σ would be related to the intertemporal labor supply elasticity because it measures the change in hours of work for a given person resulting from a particular change in his wage rate.

The statistically interesting part of the problem arises when one observes the same person for more than two periods. Suppose, for example, that we have a sample containing 1,000 workers and that each worker in our data is observed over a period of 20 years. Although one could imagine differencing the data a number of times, there exists a statistically easier procedure that effectively does the same thing. In particular, we would stack all the 20 observations for a particular worker across all workers. The new regression model, therefore, would have 20,000 observations. We would then estimate the following regression model on this stacked data set:

$$H_{it} = \sigma w_{it} + \alpha_1 F_1 + \alpha_2 F_2 + \ldots + \alpha_{1000} F_{1000} + \text{Other variables} \qquad \textbf{(2-15)}$$

where F_1 is a "dummy variable" set equal to one if that observations refers to person 1, and zero otherwise; F_2 is another dummy variable set equal to one if that observation refers to person 2, and zero otherwise; and so on. In effect, the regression model in equation (2-15) includes a dummy variable for each person in the data, and there would be 1,000 such dummy variables.

The set of dummy variables $(F_1, \ldots, F_{1000})$ are called **fixed effects,** because they indicate that hours of work for worker i, for whatever reasons, has a fixed factor that determines the person's hours of work on a permanent basis, even apart from year-to-year wage fluctuations. Put differently, the set of individual-specific fixed effects included in the regression model in equation (2-15) controls for any factors that are specific to persons and allows us to concentrate on measuring how wage changes affect changes in hours of work for a specific person. In fact, it can be shown that if each worker in our data were only observed twice, the method of including fixed effects in the regression model would be numerically identical to the common-sense differencing of all of the variables illustrated in equation (2-15).

[46] See Thomas E. MaCurdy, "An Empirical Model of Labor Supply in a Life-Cycle Setting," *Journal of Political Economy* 89 (December 1981): 1059–1085. See also Richard Blundell, Costas Meghir, and Pedro Neves, "Labour Supply and Intertemporal Substitution," *Journal of Econometrics* 59 (September 1993): 137–160; and David Card, "Intertemporal Labor Supply: An Assessment," in Christopher A. Sims, editor, *Advances in Econometrics: Sixth World Congress*, vol. II, Cambridge: Cambridge University Press, 1994, pp. 49–80.

The elasticities of intertemporal substitution estimated by the method of fixed effects tend to be positive, but numerically small. As noted above, many of the estimates suggest that the elasticity is around 0.1, indicating that a year-to-year wage increase of 10 percent would increase annual hours of work by only about 1 percent.

The statistical method of fixed effects has become a commonly used empirical technique in the toolkit of modern labor economics. It is easy to see why: There are obviously many person-specific factors that affect how many hours we work. Some of us are workaholics, and some of us would rather watch *Jersey Shore*. Our tastes for work are, to a large extent, fixed; they are a part of who we are. The individual-specific fixed effects help control for these idiosyncratic differences among workers and allow us to focus on what is most important in terms of the economic models: how changes in economic opportunities for a given worker affect the labor supply of that worker.

Labor Supply over the Business Cycle

Not only does labor supply respond to changes in economic opportunities over a worker's life cycle, but the worker also may adjust his labor supply to take advantage of changes in economic opportunities induced by business cycles. Do recessions motivate many persons to enter the labor market in order to "make up" the income of family members who have lost their jobs? Or do the unemployed give up hope of finding work in a depressed market and leave the labor force altogether?

The **added worker effect** provides one possible mechanism for a relation between the business cycle and the labor force participation rate. Under this hypothesis, so-called secondary workers who are currently out of the labor market (such as young persons or mothers with small children) are affected by the recession because the main breadwinner becomes unemployed or faces a wage cut. As a result, family income falls and the secondary workers get jobs to make up the loss. The added worker effect thus implies that the labor force participation rate of secondary workers has a countercyclical trend (that is, it moves in a direction opposite to the business cycle): it rises during recessions and falls during expansions.

The relationship between the business cycle and the labor force participation rate also can arise because of the **discouraged worker effect.** The discouraged worker effect argues that many unemployed workers find it almost impossible to find jobs during a recession and simply give up. Rather than incur the costs associated with fruitless job search activities, these workers decide to wait out the recession and drop out of the labor force. As a result of the discouraged worker effect, the labor force participation rate has a procyclical trend: it falls during recessions and increases during expansions.

Of course, the business cycle will generate both added workers and discouraged workers. The important question, therefore, is which effect dominates empirically. This question is typically addressed by correlating the labor force participation rate of a particular group with the aggregate unemployment rate, a summary measure of aggregate economic activity. If the added worker effect dominates, the correlation should be positive because the deterioration in economic conditions encourages more persons to enter the labor market. If the discouraged worker effect dominates, the correlation should be negative because the high level of unemployment in the economy convinces many workers to "give up" on their

job searches and drop out of the labor market. There is overwhelming evidence that the correlation between the labor force participation rates of many groups and the aggregate unemployment rate is negative, so the discouraged worker effect dominates.[47]

Because the discouraged worker effect dominates the correlation between labor force participation and the business cycle, the official unemployment rate reported by the Bureau of Labor Statistics (BLS) might be too low. Recall that the BLS defines the unemployment rate as the ratio of persons who are unemployed to persons who are in the labor force (that is, the employed plus the unemployed). If an unemployed person becomes discouraged and leaves the labor force, he or she is no longer actively looking for work and, hence, will no longer be counted among the unemployed. As a result, the official unemployment rate may greatly understate the unemployment problem in the aggregate economy during severe recessions. However, the argument that the discouraged workers should be included in the unemployment statistics is open to question.[48] Some of these discouraged workers may be "taking advantage" of the relatively poor labor market conditions to engage in leisure activities.

As we saw earlier, the life cycle model of labor supply suggests that some workers choose to allocate time to the labor market during certain periods of the life cycle and to consume leisure during other periods. The real wage typically rises during expansions (when the demand for labor is high) and declines during recessions (when the demand for labor slackens). We would then expect the labor force participation rate to be high at the peak of economic activity and to decline as economic conditions worsen. The procyclical trend in the labor force participation rate then arises not because workers give up hope of finding jobs during recessions but because it is not worthwhile to work in those periods when the real wage is low. In an important sense, therefore, the so-called discouraged workers are doing precisely what the life cycle model of labor supply suggests that they should do: allocate their time optimally over the life cycle by consuming leisure when it is cheap to do so. As a result, the pool of hidden unemployed should not be part of the unemployment statistics. We will discuss the implications of this controversial hypothesis in more detail in Chapter 12.

Job Loss and the Added Worker Effect

It is worth emphasizing that the business cycle is not the only economic shock that can generate added worker and discouraged worker effects. A family's economic stability—and the distribution of labor supply within the household—also will be affected by any random events that create job instability for primary earners in the household, such as unforeseen plant closings and other types of job displacement.

---

[47] Jacob Mincer, "Labor Force Participation and Unemployment: A Review of Recent Evidence," in R. A. Gordon and M. S. Gordon, editors, *Prosperity and Unemployment,* New York: Wiley, 1966, pp. 73–112; and Shelly Lundberg, "The Added Worker Effect," *Journal of Labor Economics* 3 (January 1985): 11–37.

[48] This argument is developed at length in the influential article by Robert E. Lucas and Leonard Rapping, "Real Wages, Employment, and Inflation," *Journal of Political Economy* 77 (October 1969): 721–754.

Most of us look forward to days when we can take off some time from work and just relax. In an ideal world, on such days the weather would cooperate: It would be sunny and warm, and we would be totally free to enjoy our favorite activities, whether they be playing volleyball on the beach, sitting down in our backyard, or strolling down a shop-filled avenue.

Unfortunately, the weather does not always cooperate. As Mark Twain famously said, "Everybody talks about the weather, but nobody does anything about it." Well, it turns out that although our actions cannot affect the weather, we *can* react to the weather in our area and take actions that minimize the adverse impact that bad weather would have had on our planned leisure activities.

The life cycle version of the neoclassical labor-leisure model predicts that individuals work more in those periods when the rewards to working are the greatest. If the weather interferes with leisure activities—for example, making a day at the beach much less pleasant—then it would be optimal for a person to work more on that day and postpone the day at the beach until a day when the sun cooperates.

It turns out that working men do adjust their work-leisure activities in accordance with this very intuitive prediction of the life cycle model. Suppose a "rainy day"

occurs when there is precipitation of at least 0.1 inch in the local area. Holding other factors constant, including a person's education, age, and average weather patterns in the region of the country where the worker lives, a rainy day increases the time allocated to working activities by 29 minutes per day, while it reduces the time allocated to leisure activities by 25 minutes per day. In other words, the typical male worker will adjust his time over the working week to account for local weather fluctuations, working more on those days when the weather makes leisure activities less pleasant and consuming more leisure activities on those days when it is worth consuming.

Surprisingly, these very intuitive empirical patterns are not found among working women. One potential explanation of the gender difference may be that men are more likely to participate in sports activities during their leisure hours. But it is unclear if this conjecture accounts for the entire gender difference.

Sources: Marie Connolly, "Here Comes the Rain Again: Weather and the Intertemporal Substitution of Leisure," *Journal of labor Economics* 24 (January 2008): 73–100; see also Jorge Gonzalez-Chapela, "On the Price of Recreation Goods as a Determinant of Male Labor Supply," *Journal of Labor Economics* 25 (October 2007): 795–824.

Recent research shows that intra-household responses in labor supply play an important role in attenuating earnings losses caused by layoffs and plant closings. It is documented, for instance, that there is a sizable positive labor supply response by the wife to the husband's unexpected job loss, and that this supply increase can compensate for over 25 percent of the loss in family income.[49] Interestingly, the evidence also indicates that much of the potential increase in the wife's labor supply will be "crowded out" by the presence of the unemployment insurance system. In other words, the government-provided assistance to the unemployed husband greatly reduces the need for the wife to enter the labor market in response to the husband's job loss.

[49] Melvin Stephens, "Worker Displacement and the Added Worker Effect," *Journal of Labor Economics* 20 (July 2002): 504–537; see also Julie Berry Cullen and Jonathan Gruber, "Does Unemployment Insurance Crowd Out Spousal Labor Supply," *Journal of Labor Economics* 18 (July 2000): 546–572.

2-13 Policy Application: The Decline in Work Attachment among Older Workers

As noted earlier, there has been a marked drop in labor force participation among older men. It is hard to argue that the increasing propensity for early retirement is linked to the deteriorating health of this particular age group. After all, at the same time that their labor market attachment was weakening, the life expectancy of white men aged 50 rose from 22 to 29.2 years between 1939 and 2007.[50]

Part of the declining labor force participation of older workers may be attributable to an increase in pension benefits. The fraction of men who were covered by pension programs other than Social Security rose from 26 percent in 1950 to 66 percent in 1990. Not surprisingly, there is a strong link between the availability of private pension plans and the labor force participation of older men. For example, the probability that men aged 58 to 63 work falls by 18 percentage points if they have private pension plans.[51]

Many studies have attempted to determine if the increased generosity of the Social Security system is also partly responsible for the move toward early retirement. After accounting for inflation, Social Security benefits increased by about 20 percent during the early 1970s. Moreover, during the 1980s, a period when real wages fell for many workers, real Social Security benefits (which are indexed to the inflation rate) remained roughly constant. Despite the substantial increase in a worker's "Social Security wealth" (or the total value of the Social Security benefits that the worker can expect to receive over his lifetime), the available evidence does not strongly support the argument that increases in Social Security benefits explain a large part of the decline in the participation rates of older persons. In fact, the evidence suggests that at most 15 percent of the decline in participation rates of older workers can be attributed to the increase in Social Security retirement benefits.[52]

Some studies instead argue that an important part of the decline in the labor market attachment of older workers in the United States can be attributed to the work disincentives created by the Social Security Disability Program. In the United States, workers who become disabled are eligible to receive disability payments for as long as the disability lasts. The monthly disability benefit equals the Social Security retirement benefits that the worker would have received had he or she continued working until age 65, *regardless of the worker's age at the time the disability occurred.*

---

[50] U.S. Bureau of the Census, *Statistical Abstract of the United States,* Washington, DC: Government Printing Office, various issues.

[51] Alan Gustman and Thomas Steinmeier, "Partial Retirement and the Analysis of Retirement Behavior," *Industrial and Labor Relations Review* 37 (April 1984): 403–415; and Edward P. Lazear, "Pensions as Severance Pay," in Zvi Bodie and John Shoven, editors, *Financial Aspects of the United States Pension System,* Chicago: University of Chicago Press, 1983.

[52] Alan B. Krueger and Jörn-Steffen Pischke, "The Effect of Social Security on Labor Supply: A Cohort Analysis of the Notch Generation," *Journal of Labor Economics* 10 (October 1992): 412–437. See also Gary Burtless, "Social Security, Unanticipated Benefit Increases, and the Timing of Retirement," *Review of Economic Studies* 53 (October 1986): 781–805.

Audrey was born in March 1916 and her sister Edith was born in June 1917. They both began working at the same book bindery in southern California in October 1957. They both worked continuously at this firm and received the same pay until they retired. When the younger sister Edith turned 65, both Edith and Audrey went to the Social Security office to claim their benefits. Because Audrey had worked for about 18 months past her 65th birthday, she expected to receive slightly higher benefits. It turned out, however, that Audrey received $624.40 per month, whereas Edith received only $512.60 per month.

This real-life example illustrates the decline in economic opportunities experienced by the so-called notch babies, the cohort of persons born between 1917 and 1921, in their retirement years. Because of a 1977 legislative change in the formulas used to calculate Social Security benefits, the notch cohort received substantially lower benefits than earlier cohorts. As the experience of Audrey and Edith illustrates, a worker born in

1917 can receive about 20 percent less Social Security income than a worker born in 1916 with essentially the same job and earnings history.

The hypothesis that an increase in Social Security benefits reduces labor force participation rates must imply that a substantial decrease in benefits (like the one experienced by the notch babies) should increase labor force participation rates. It turns out, however, that the labor force participation rate of the notch babies is not markedly higher than the participation rate of other birth cohorts. The "natural experiment" arising from the legislative creation of the notch babies, therefore, suggests that increases in Social Security wealth can only explain a minor part of the decline in the labor force participation rates of older workers.

Sources: Alan B. Krueger and Jörn-Steffen Pischke, "The Effect of Social Security on Labor Supply: A Cohort Analysis of the Notch Generation," *Journal of Labor Economics* 10 (October 1992): 412–437.

Many workers would like to claim that they are disabled in order to enjoy the leisure activities associated with early retirement. As a result, the eligibility requirements for the disability program are harsh and strictly enforced. Workers applying for disability benefits must often be certified as being disabled by government-picked physicians; there is a waiting period of five months before the worker can apply for disability benefits; and the worker cannot be employed in "gainful activities" (defined as a job where the worker earns more than $500 per month).

There is heated disagreement over whether the disability program has contributed to the decline in the labor force participation of older workers. Some studies have claimed that practically the entire decline in labor force participation rates among men aged 55 to 64 can be attributed to the disability program.[53] Other researchers, however, cast doubt

[53] Donald Parsons, "The Decline in Male Labor Force Participation," *Journal of Political Economy* 88 (February 1980): 117–134. See also John Bound and Timothy Waidmann, "Disability Transfers, Self-Reported Health, and the Labor Force Attachment of Older Men: Evidence from the Historical Record," *Quarterly Journal of Economics* 107 (November 1992): 1393–1420; David H. Autor and Mark G. Duggan, "The Rise in the Disability Rolls and the Decline in Unemployment," *Quarterly Journal of Economics* 118 (February 2003): 157–205; and Dan Black, Kermit Daniel, and Seth Sanders, "The Impact of Economic Conditions on Participation in Disability Programs: Evidence from the Coal Boom and Bust," *American Economic Review* 92 (March 2002): 27–50.

on these findings. One recent study, for example, examined the labor supply decisions of the disability applicants who are rejected by the government.[54] Because of the strict eligibility requirements, the government rejects nearly half of the claims. If these rejected claims were mainly attempts by workers to misuse the program, one might expect that the rejected workers would return to the labor force once they learned that they cannot "get away" with this early retirement strategy. It turns out, however, that fewer than half of the rejected applicants go back to work after the final (and adverse) determination of their case. This result has been interpreted as indicating that the men who receive disability benefits would not have been in the labor force even in the absence of such a program.

Despite these criticisms, there remains a strong suspicion that the disability program has much to do with the increase in early retirement. Perhaps the most convincing evidence is provided by a recent study of the Canadian experience.[55] In the United States, the disability program is a federal program, which implies that eligibility and benefit levels are the same throughout the entire country. In Canada, there are two programs: the Quebec Pension Program (QPP) covers only persons residing in Quebec and the Canada Pension Program (CPP) covers persons residing in the rest of Canada. Although these two systems are similar in many ways, benefits in the QPP rose faster in the 1970s and 1980s. By 1986, the QPP was substantially more generous than the CPP. In January 1987, the CPP raised its benefit levels to bring the two programs to parity.

Table 2-6 provides a difference-in-differences analysis of the impact of this change in benefit levels on the labor supply of the affected population. The top rows of the table show that benefit levels in the rest of Canada increased by $2,642 (Canadian dollars) between

TABLE 2-6   The Impact of Disability Benefits on Labor Supply in Canada

Source: Jonathan Gruber, "Disability Insurance Benefits and Labor Supply," *Journal of Political Economy* 108 (December 2000): 1175.

	Before	After	Difference	Difference-in-Differences
Annual benefits:				
Canada Pension Program	$5,134	$7,776	$2,642	$1,666
Quebec Pension Program	6,876	7,852	976	
Percent of men aged 45–59 not employed last week:				
Treatment group: CPP	20.0%	21.7%	1.7%	2.7%
Control group: QPP	25.6	24.6	−1.0	

[54] John Bound, "The Health and Earnings of Rejected Disability Insurance Applicants," *American Economic Review* 79 (June 1989): 482–503.

[55] Jonathan Gruber, "Disability Insurance Benefits and Labor Supply," *Journal of Political Economy* 108 (December 2000): 1162–1183.

1986 and 1987, as compared to only an increase of $976 in Quebec. As a result of the policy shift, the average disability benefit in the rest of Canada increased by $1,666 more than the increase experienced by persons residing in Quebec.

The bottom rows of the table document how this increased generosity affected labor supply. The fraction of men aged 45–59 who did not work fell from 25.6 to 24.6 in Quebec (a decrease of 1.0 percentage point), likely reflecting changes in aggregate economic activity over the period. In contrast, the proportion of comparable men residing outside Quebec who did not work *rose* from 20.0 to 21.7 percent, an increase of 1.7 percentage points. The difference-in-differences estimator (or $1.7 - (-1.0)$) implies that the increased generosity of the disability program increased the proportion of men who did not work by 2.7 percentage points. It seems, therefore, that generous disability benefits do indeed reduce the labor supply of men nearing retirement age.

The Social Security Earnings Test

Many workers who consider themselves retired continue to work, perhaps in a part-time job. In the United States, for example, nearly 20 percent of "retired" persons also hold a job.

Until 2000, the Social Security system had a provision, known as the **Social Security earnings test,** that presumably discouraged Social Security recipients from working. In the year 2000, for example, retirees between the ages of 65 and 69 who received Social Security benefits could have earned up to $17,000 per year without affecting their retirement benefits.[56] If earnings exceeded this threshold, the government reduced the size of the Social Security benefit. In particular, $1 of Social Security benefits was withheld for every $3 earned *above* the exempt amount, so that workers who earned more than $17,000 implicitly faced a 33 percent tax rate. The earnings test did not apply to workers who were 70 or older. In 2000, the earnings test was eliminated and retired workers are now free to work and collect Social Security benefits without any penalty on their benefits.

It was often claimed that the earnings test discouraged retirees from participating in the labor force. It turns out, however, that these claims were not justified. Figure 2-22 shows how the earnings test could affect work incentives. Suppose that the retiree receives $10,000 in Social Security benefits per year (and that he does not have any other nonlabor income). Let us now construct the budget line facing this worker under the Social Security system in effect in the year 2000. The endowment point E in the figure indicates that if the retiree does not work, he could purchase $10,000 worth of goods. If the retiree works a few hours (at a wage of w dollars), he can increase the value of his consumption bundle, as illustrated by the segment FE of the budget line.

At point F in the figure, the retiree earns the maximum allowed by the Social Security Administration before Social Security benefits are reduced, so he can consume $27,000 worth of goods (the $10,000 Social Security benefits plus $17,000 in labor earnings). If the retiree keeps on working, however, the marginal wage rate is no longer w, but $w(1 - 0.33)$,

[56] See Leora Friedberg, "The Labor Supply Effects of the Social Security Earnings Test," *Review of Economics and Statistics* 82 (February 2000): 48–63; and Steven J. Haider and David S. Loughran, "The Effect of the Social Security Earnings Test on Male Labor Supply: New Evidence from Survey and Administrative Data," *Journal of Human Resources* 43 (Winter 2008): 57–87.

FIGURE 2-22 The Impact of the Social Security Earnings Test on Hours of Work

The Social Security earnings test (which taxes retirees when they earn more than $17,000 per year) generates the budget "line" *HGFE*. The repeal of the earnings test moves retirees to budget line *H'E*. The first retiree (worker 1) would not change his hours of work; the second retiree would reduce his hours; and the third retiree might increase or decrease his hours, depending on whether substitution or income effects dominate.
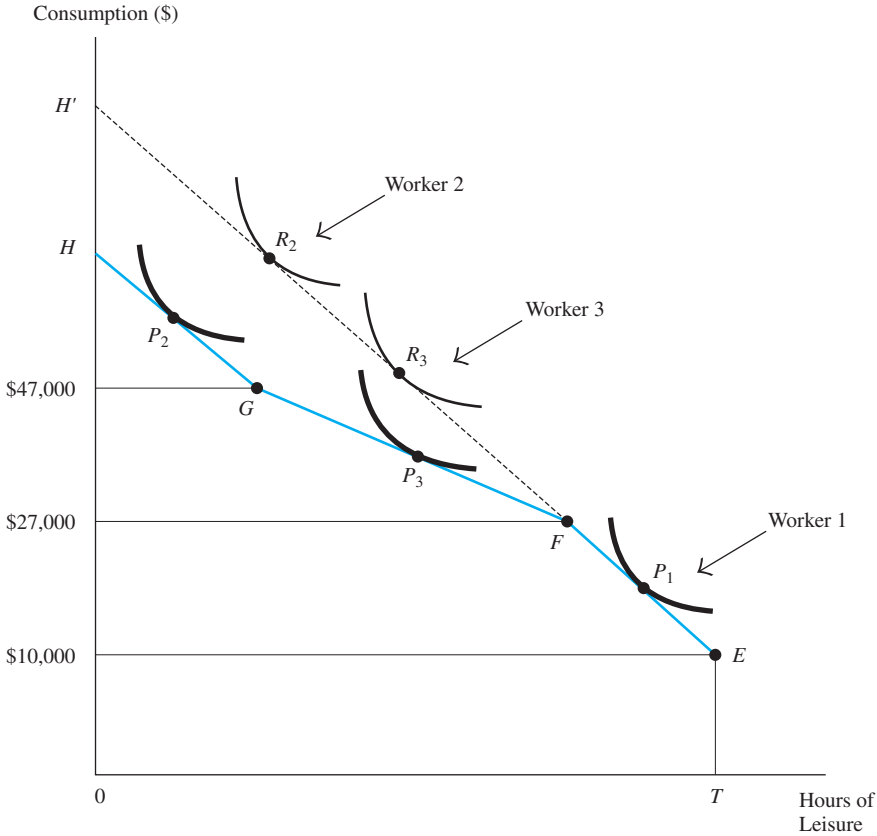
flattening out the budget line, and generating segment *FG*. Finally, if the retiree earns more than $47,000, the retiree forfeits his entire Social Security benefits, and the slope of the budget line reverts back to its original slope.[57] The earnings test thus generates the budget "line" *HGFE* in Figure 2-22.

It is of interest to ask if the elimination of the earnings test would increase the labor supply of older workers. The elimination of the test would allow the retiree to face budget line *H'E*, instead of *HGFE*. As is evident in Figure 2-22, there are three potential effects of the

[57] The first $17,000 of earnings for this retiree is exempt from the Social Security tax, so that only $30,000 of wage income is subject to the tax. Because Social Security benefits are reduced by $1 for every $3 of taxable income, the entire Social Security benefit of a worker who earns $47,000 is taxed away. The consumption basket available to this worker is illustrated by point *G* in Figure 2-22. He has $47,000 available for consumption (or $10,000 in Social Security benefits + $47,000 in wage income − $10,000 in Social Security taxes).

earnings test on work incentives. The first worker (worker 1 in the figure) has indifference curves that place him at point P_1, where he works very few hours, regardless of whether there is an earnings test. Obviously, this retiree will not be affected by the elimination of the earnings test. The second worker's indifference curves place him at point P_2, indicating strong "tastes for work." This person allocates many hours to the labor market even though it means he has to forfeit his Social Security benefits. Interestingly, removing the earnings test for this worker is equivalent to an increase in the person's wealth, moving the worker from point P_2 to point R_2. This income effect induces the retiree to consume more leisure hours, thus *reducing* work hours.

Finally, the third worker is a retiree who works a "medium" number of hours. This person has not entirely forfeited his Social Security benefits and faces a 33 percent tax rate on labor earnings. The repeal of the earnings test moves this worker from point P_3 to point R_3. In other words, this worker effectively gets a wage increase when the earnings test is repealed. As such, the worker will experience both an income and a substitution effect. The income effect will motivate the worker to consume more leisure and work fewer hours; the substitution effect induces the worker to consume fewer leisure hours and work more hours. As drawn, the substitution effect dominates.

Overall, the theory suggests that the elimination of the Social Security earnings test is unlikely to substantially increase labor supply among retirees. A few studies have examined the labor supply consequences of repealing the earnings test. The evidence confirms the theoretical expectation: the labor supply effects of the repeal tended to be small.[58]

Summary

- The reservation wage is the wage that makes a person indifferent between working and not working. A person enters the labor market when the market wage rate exceeds the reservation wage.

- Utility-maximizing workers allocate their time so that the last dollar spent on leisure activities yields the same utility as the last dollar spent on goods.

- An increase in nonlabor income reduces hours of work of workers.

- An increase in the wage generates both an income and a substitution effect among persons who work. The income effect reduces hours of work; the substitution effect increases hours of work. The labor supply curve, therefore, is upward sloping if substitution effects dominate and downward sloping if income effects dominate.

- An increase in nonlabor income reduces the likelihood that a person enters the labor force. An increase in the wage increases the likelihood that a person enters the labor force.

- The labor supply elasticity is on the order of -0.1 for men and $+0.2$ for women.

- Welfare programs create work disincentives because they provide cash grants to participants as well as tax those recipients who enter the labor market. In contrast, credits on earned income create work incentives and draw many nonworkers into the labor force.

[58] Jae G. Song and Joyce Manchester, "New Evidence on Earnings and Benefit Claims Following Changes in the Retirement Earnings Test in 2000," *Journal of Public Economics* 91 (April 2007): 669–700.

Key Concepts

Review Questions

1. What happens to the reservation wage if nonlabor income increases, and why?
2. What economic factors determine whether a person participates in the labor force?
3. How does a typical worker decide how many hours to allocate to the labor market?
4. What happens to hours of work when nonlabor income decreases?
5. What happens to hours of work when the wage rate falls? Decompose the change in hours of work into income and substitution effects.
6. What happens to the probability that a particular person works when the wage rises? Does such a wage increase generate an income effect?
7. Why do welfare programs create work disincentives?
8. Why does the earned income tax credit increase the labor force participation rate of targeted groups?
9. Why have average hours worked per week declined?
10. Why did the labor force participation rate of women increase so much in the past century?
11. Why does a worker allocate his or her time over the life cycle so as to work more hours in those periods when the wage is highest? Why does the worker not experience an income effect during those periods?
12. What is the added worker effect? What is the discouraged worker effect?
13. What factors account for the secular decline in labor force participation rates among older workers in the United States?

Problems

2-1. How many hours will a person allocate to leisure activities if her indifference curves between consumption and goods are concave to the origin?
2-2. What is the effect of an increase in the price of market goods on a worker's reservation wage, probability of entering the labor force, and hours of work?
2-3. Tom earns $15 per hour for up to 40 hours of work each week. He is paid $30 per hour for every hour in excess of 40. Tom faces a 20 percent tax rate and pays $4 per hour in child care expenses for each hour he works. Tom receives $80 in child support payments each week. There are 168 hours in the week. Graph Tom's weekly budget line.
2-4. Cindy gains utility from consumption C and leisure L. The most leisure she can consume in any given week is 168 hours. Her utility function is $U(C, L) = C \times L$. This functional form implies that Cindy's marginal rate of substitution is C/L. Cindy receives $630 each week from her great-grandmother—regardless of how much Cindy works. What is Cindy's reservation wage?

2-5. You can either take a bus or drive your car to work. A bus pass costs $5 per week, whereas driving your car to work costs $60 weekly (parking, tolls, gas, etc.). You spend half an hour less on a one-way trip in your car than on a bus. How would you prefer to travel to work if your wage rate is $10 per hour? Will you change your preferred mode of transportation if your wage rate rises to $20 per hour? Assume you work five days a week and time spent riding on a bus or driving a car does not directly enter your utility.

2-6. Shelly's preferences for consumption and leisure can be expressed as

$$U(C, L) = (C - 200) \times (L - 80)$$

This utility function implies that Shelly's marginal utility of leisure is $C - 200$ and her marginal utility of consumption is $L - 80$. There are 168 hours in the week available to split between work and leisure. Shelly earns $5 per hour after taxes. She also receives $320 worth of welfare benefits each week regardless of how much she works.

a. Graph Shelly's budget line.

b. What is Shelly's marginal rate of substitution when L = 100 and she is on her budget line?

c. What is Shelly's reservation wage?

d. Find Shelly's optimal amount of consumption and leisure.

2-7. Explain why a lump-sum government transfer can entice some workers to stop working (and entices no one to start working) while the earned income tax credit can entice some people who otherwise would not work to start working (and entices no one to stop working).

2-8. In 1999, 4,860 TANF recipients were asked how many hours they worked in the previous week. In 2000, 4,392 of these recipients were again subject to the same TANF rules and were again asked their hours of work during the previous week. The remaining 468 individuals were randomly assigned to a "Negative Income Tax" (NIT) experiment that gave out financial incentives for welfare recipients to work and subjected them to its rules. Like the other group, they were asked about their hours of work during the previous week. The data from the experiment are contained in the table below.

a. What effect did the NIT experiment have on the employment rate of public assistance recipients? Develop a standard difference-in-differences table to support your answer.

b. What effect did the NIT experiment have on the weekly hours worked of public assistance recipients who worked positive hours during the survey week? Develop a standard difference-in-differences table to support your answer.

	Total Number of Recipients	Number of Recipients Who Worked at Some Time in the Survey Week		Total Hours of Work by All Recipients in the Survey Week	
		1999	2000	1999	2000
TANF	4,392	1,217	1,568	15,578	20,698
NIT	468	131	213	1,638	2,535
Total	4,860	1,348	1,781	17,216	23,233

2-9. Consider two workers with identical preferences, Phil and Bill. Both workers have the same life cycle wage path in that they face the same wage at every age, and they know what their future wages will be. Leisure and consumption are both normal goods.

    a. Compare the life cycle path of hours of work between the two workers if Bill receives a one-time, unexpected inheritance at the age of 35.

    b. Compare the life cycle path of hours of work between the two workers if Bill had always known he would receive (and, in fact, does receive) a one-time inheritance at the age of 35.

2-10. Under current law, most Social Security recipients do not pay federal or state income taxes on their Social Security benefits. Suppose the government proposes to tax these benefits at the same rate as other types of income. What is the impact of the proposed tax on the optimal retirement age?

2-11. A worker plans to retire at the age of 65, at which time he will start collecting his retirement benefits. Then there is a sudden change in the forecast of inflation when the worker is 63 years old. In particular, inflation is now predicted to be higher than it had been expected so that the average price level of market goods and wages is now expected to be higher. What effect does this announcement have on the person's preferred retirement age

    a. If retirement benefits are fully adjusted for inflation?

    b. If retirement benefits are not fully adjusted for inflation?

2-12. Presently, there is a minimum and maximum social security benefit paid to retirees. Between these two bounds, a retiree's benefit level depends on how much she contributed to the system over her work life. Suppose Social Security was changed so that everyone aged 65 or older was paid $12,000 per year regardless of how much she earned over her working life or whether she continued to work after the age of 65. How would this likely affect hours worked of retirees?

2-13. Over the last 100 years, real household income and standards of living have increased substantially in the United States. At the same time, the total fertility rate, the average number of children born to a woman during her lifetime, has fallen in the United States from about three children per woman in the early twentieth century to about two children per woman in the early twenty-first century. Does this suggest that children are inferior goods?

2-14. Consider a person who can work up to 80 hours each week at a pretax wage of $20 per hour but faces a constant 20 percent payroll tax. Under these conditions, the worker maximizes her utility by choosing to work 50 hours each week. The government proposes a negative income tax whereby everyone is given $300 each week and anyone can supplement her income further by working. To pay for the negative income tax, the payroll tax rate will be increased to 50 percent.

    a. On a single graph, draw the worker's original budget line and her budget line under the negative income tax.

    b. Show that the worker will choose to work fewer hours if the negative income tax is adopted.

    c. Will the worker's utility be greater under the negative income tax?

2-15. The absolute value of the slope of the consumption-leisure budget line is the after-tax wage, *w*. Suppose some workers earn *w* for up to 40 hours of work each week and then earn 2*w* for any hours worked thereafter (called overtime). Other workers may earn *w* for up to 40 hours of work each week and then only earn 0.5*w* thereafter as working more than 40 hours requires getting a second job, which pays an hourly wage less than their primary job. Both types of workers experience a "kink" in their consumption-leisure budget line.

 a. Graph in general terms the budget line for each type of worker.

 b. Which type of worker is likely to work up to the point of the kink, and which type of worker is likely to choose a consumption-leisure bundle far away from the kink?

Selected Readings

David H. Autor and Mark G. Duggan, "The Rise in the Disability Rolls and the Decline in Unemployment," *Quarterly Journal of Economics* 118 (February 2003): 157–205.

Gary S. Becker, "A Theory of the Allocation of Time," *Economic Journal* 75 (September 1965): 493–517.

Stacy Dickert-Conlin and Amitabh Chandra, "Taxes and the Timing of Births," *Journal of Political Economy* 107 (February 1999): 161–177.

Nada Eissa and Jeffrey B. Liebman, "Labor Supply Response to the Earned Income Tax Credit," *Quarterly Journal of Economics* 111 (May 1996): 605–637.

Jeffrey Grogger and Charles Michalopoulos, "Welfare Dynamics under Time Limits," *Journal of Political Economy* 111 (June 2003): 530–554.

James J. Heckman, "Life Cycle Consumption and Labor Supply: An Explanation of the Relationship between Income and Consumption over the Life Cycle," *American Economic Review* 64 (March 1974): 188–194.

James J. Heckman, "Sample Selection Bias as a Specification Error with an Application to the Estimation of Labor Supply Functions," in James P. Smith, editor, *Female Labor Supply: Theory and Estimation.* Princeton, NJ: Princeton University Press, 1980, pp. 206–248.

Thomas E. MaCurdy, "An Empirical Model of Labor Supply in a Life-Cycle Setting," *Journal of Political Economy* 89 (December 1981): 1059–1085.

Jacob Mincer, "Labor Force Participation of Married Women," in H. Gregg Lewis, editor, *Aspects of Labor Economics.* Princeton, NJ: Princeton University Press, 1962, pp. 63–97.

Robert A. Moffitt, "Welfare Programs and Labor Supply," in Alan J. Auerbach and Martin Feldstein, editors, *Handbook of Public Economics,* vol. 4. Amsterdam: Elsevier, 2006.

Web Links

The Bureau of Labor Statistcs publishes a detailed description of how it defines and measures the concepts of labor force and unemployment: stats.bls.gov/cps/cps_htgm.htm.

The Social Security Administration publishes many documents that provide not only a detailed description of the system, but also such facts as the most popular names given to babies born in a particular calendar year and a calculator that predicts Social Security benefits for a particular worker: www.ssa.gov/.