CHAPTER 5

FOSTERING PRODUCTIVITY GROWTH

In 1870, a family farmer planting corn in Iowa would have expected to grow 35 bushels an acre. Today, that settler's descendant can grow nearly 180 bushels an acre and uses sophisticated equipment to work many times the acreage of his or her forbearer. Because of higher yields and the use of time-saving machinery, the quantity of corn produced by an hour of farm labor has risen from an estimated 0.64 bushel in 1870 to more than 60 bushels in 2013. This 90-fold increase in labor productivity—that is, bushels of corn (real output) an hour—corresponds to an annual rate of increase of 3.2 percent compounded over 143 years. In 1870, a bushel of corn sold for approximately $0.80, about two days of earnings for a typical manufacturing worker; today, that bushel sells for approximately $4.30, or 12 minutes worth of average earnings.[1]

This extraordinary increase in corn output, fall in the real price of corn, and the resulting improvement in physical well-being, did not come about because we are stronger, harder-working, or tougher today than the early settlers who first plowed the prairies. Rather, through a combination of invention, more advanced equipment, and better education, the Iowa farmer today uses more productive strains of corn and sophisticated farming methods to get more output an acre. Today's farmer harnesses more capital equipment, such as advanced planters and combines, to plant more acres, and has the know-how to operate this sophisticated equipment.

Technological advances such as corn hybridization, fertilizer technology, disease resistance, and mechanical planting and harvesting have resulted from decades of research and development. While the government has supported some of this research and its dissemination—for example, through basic biological research and land-grant universities—much of this research occurred in the private sector. However, the government has

[1] Sources: Parker and Klein (1966), 1870 *Census of Manufacturers*, Iowa State University Extension Service (2013), Bureau of Labor Statistics, USDA *Economic Research Service*.

facilitated this private-sector technological innovation by providing the infrastructure to transport and sell increasing quantities of the products and a regulatory and legal environment, such as the U.S. patent system, which clarifies and enforces rights to inventions (more generally, to intellectual property) so that the private sector can reap the rewards of research. These property rights create incentives for innovators, while also allowing others to build on their inventions. The improvements in productivity made possible by technological progress have appeared not just in agriculture, but also throughout the U.S. economy.

The framework of government support for technological innovation is facing new challenges that stem from an ever-changing scientific and legal landscape. Many of these challenges center on the best way to support and encourage development of intellectual property which now encompasses improvements, not just to tractor design, but also technological changes to the software that optimizes its performance. Farmers can now use the Internet to do market research, purchase inputs, make direct sales, and participate in online crop and livestock auctions. Other challenges involve issues surrounding the allocation of the electromagnetic spectrum in a way that supports the efficient development of new wireless and communications technologies that will improve productivity and connectivity—for the farmer in the combine's cab as well as for millions of other consumers and businesses—while weighing national security and other concerns. These challenges also include striking the appropriate balance between the need for the government to support fundamental research, which can have large positive externalities that will not be realized by any individual private actor, and the importance of private-sector innovation in driving technology forward.

Another set of challenges relates to how the gains from innovation are shared. In the decades following World War II, productivity improvements translated relatively automatically into compensation increases for families across the income spectrum. But starting in the 1970s, inequality began its relentless rise and productivity growth became increasingly disconnected from compensation growth for typical families. The trends in inequality are related to the trends in productivity, as well as to other broad economic trends. Some of the technological changes over the past three decades, especially those related to information technology, have raised the relative reward to skills obtained through advanced academic study. Thus, the slowing growth of educational attainment both potentially slows innovation and increases inequality by raising the returns to the most highly educated workers. Although expanding the size of markets through globalization can help increase the productivity of the economy, it can also create challenges for inequality.

This chapter begins with a review of the history of productivity growth since World War II, emerging inequality trends, and the government's role in fostering productivity growth. It then focuses on two important current issues in more detail: wired and wireless broadband infrastructure and the efficient allocation of the electromagnetic spectrum; and, new challenges to the U.S. patent system posed by standard-essential patents and patent-assertion entities.

Trends in Total Factor Productivity

The most commonly used measure of productivity is *labor productivity*—that is, real output per hour worked. Over the long run, improvements in labor productivity translate into growth of output, wages, and income. Labor productivity can grow for multiple reasons: more capital per worker (increased capital intensity), increased labor skills (a more experienced workforce, more and better education and training), and technological advances that improve the quality and productivity for a given level of capital and labor skills (inventions, technological progress, process improvements, and other factors).

Because of the importance of technological progress in enhancing long-run growth, economists also use another measure of productivity called *total factor productivity*, or TFP, which proxies for the effect of technological progress. From 1948 to 2012, labor productivity growth in the private nonfarm business sector has averaged 2.2 percent per year, and total factor productivity growth, as measured by the series on multifactor productivity produced by the Bureau of Labor Statistics (BLS), has averaged 1.1 percent per year. This growth of productivity has not been constant, however, and can usefully be thought of as occurring in three episodes: a period of fast productivity growth through the early 1970s, a period of slow productivity growth through the mid-1990s, and a period of somewhat faster productivity growth since then, but still not as fast as in the 1950s and 1960s.

Labor Productivity, Total Factor Productivity, and Multifactor Productivity

The growth rate of labor productivity equals the growth rate of output, minus the growth rate of labor input (worker hours), thus yielding the growth rate of output per worker hour. In contrast, the growth rate of TFP is the growth rate of output, minus the growth rate of output that would be expected solely from the growth rate of the inputs to production. The resulting gap between the actual growth rate of output and the growth rate arising solely because of the growth of inputs is also known as the Solow residual, and is a measure of how well those inputs are combined. Thus, the growth rate of TFP tracks a broadly defined concept of technological change that encompasses scientific innovation and invention, managerial innovations, effects of reorganization of the production process, and other efficiency improvements that do not accrue uniquely to a single measured input.

The concept of total factor productivity is appealing because it estimates the contribution of technological developments to economic growth, and because it can be applied at the level of an industry as well as to the overall economy. In practice, measuring TFP poses several challenges. First, TFP is not observed directly and instead must be estimated using measured inputs and estimates of how the inputs contribute to output. Second, the inputs discussed so far have been capital and labor, but other inputs to production also include, in particular, energy, materials, and business services. Third, for a given level of other inputs, output can increase by hiring better-trained or higher-skilled workers; so for the purpose of measuring TFP, the desired concept of labor input captures changes in both the quantity and quality of labor input. Because labor quality is not observed, proxies such as age and education must be used. Both academics and the U.S. Government have tackled these and other measurement challenges, and have developed estimates of the growth of TFP. This chapter uses an estimate of TFP produced by the Bureau of Labor Statistics called *multifactor productivity*, or MFP, which is described in Box 5-1.[2]

Postwar U.S. Productivity Growth

According to the BLS measure of labor productivity shown in Table 5-1, an American worker could produce more than four times as much output per hour in 2012 as in 1948.[3] Because MFP takes into account the

[2] One of the many other challenges in estimating total factor productivity is that the intensity of utilization of inputs varies over the business cycle. For example, because hiring and training workers is expensive, firms might retain some workers in a mild downturn, so that fluctuations in output are greater than fluctuations in employment (a relationship which, when recast in terms of the unemployment rate, is known as Okun's Law). The BLS MFP series does not adjust for changes in factor utilization, which can produce cyclical fluctuations in MFP. Basu, Fernald, and Kimball (2006) provide an approach to adjusting for such cyclical variation, and a quarterly TFP series produced using their method is currently maintained by the Federal Reserve Bank of San Francisco (Fernald 2012).

[3] This discussion of postwar productivity performance cites statistics for nonfarm private businesses. Recall the earlier discussion about how productivity growth in farming allowed fewer resources to be devoted to it. By 1947, farming accounted for less than a nine percent share of GDP. Today that share is about one percent.

Box 5-1: Measuring Multifactor Productivity

The Bureau of Labor Statistics (BLS) publishes annual data on multifactor productivity, covering the private business sector, the private nonfarm business sector, the manufacturing sector, and 18 industries within the manufacturing sector.

Private business-sector output is a chain-type, annual-weighted (Fisher-Ideal) index constructed after excluding general government, nonprofit institutions, private households (including owner-occupied housing), and government enterprises from gross domestic product (GDP). The input measure is an aggregation of two inputs, labor and capital. Labor input is obtained by chained Tornqvist aggregation of the hours worked in private business by all persons, classified by age, education, and gender with weights determined by each group's share of total labor compensation. Capital inputs are measured based on the flow of services derived from physical assets. For each of 60 industries in the private-business sector, quantities of each capital asset are aggregated into a Tornqvist index, using estimated rental prices. Current-dollar capital costs are found by multiplying the rental price for each asset by the asset's constant-dollar stock, adjusting for capital composition effects. Finally, the combined input (labor and capital) measure is constructed via another Tornqvist index, taking as weights each inputs' share of total costs derived from the National Income and Product Accounts.

Manufacturing is treated somewhat differently. The output measure, known as sectoral output, is the value of production shipped to purchasers outside the domestic industry, either to satisfy final demand or to use as an input in other industries. Because additional inputs to manufacturing can be tracked, the input measures available include not just capital and labor, but also energy, non-energy materials, and purchased business services input. Intra-industry purchases are removed to avoid double counting. The resulting aggregate input is referred to by the acronym KLEMS—capital (K), labor (L), energy (E), materials (M) and services (S). Given these inputs and outputs, multifactor productivity is computed for 18 3-digit and 86 4-digit North American Industry Classification System (NAICS) manufacturing industries and for the manufacturing sector as a whole using the Tornqvist aggregation methods described above for the private business-sector manufacturing.

growth of capital and other factors, labor productivity growth generally exceeds MFP growth. For example, even absent technological change, labor can be more productive simply by using more capital; that is, by increasing the capital-labor ratio or so-called capital deepening. Mathematically, the growth rate of labor productivity is the sum of the MFP growth rate, the

Table 5–1
Sources of Productivity Improvement, Nonfarm Private Business, 1948–2012

Source	Improvement (multiple)	Contribution to Labor Productivity Growth (percent)
Composition of Labor	1.15	10
Capital	1.74	38
MFP	2.10	52
Labor Productivity	4.21	100

Source: Bureau of Labor Statistics, Productivity and Costs, Multifactor Productivity.

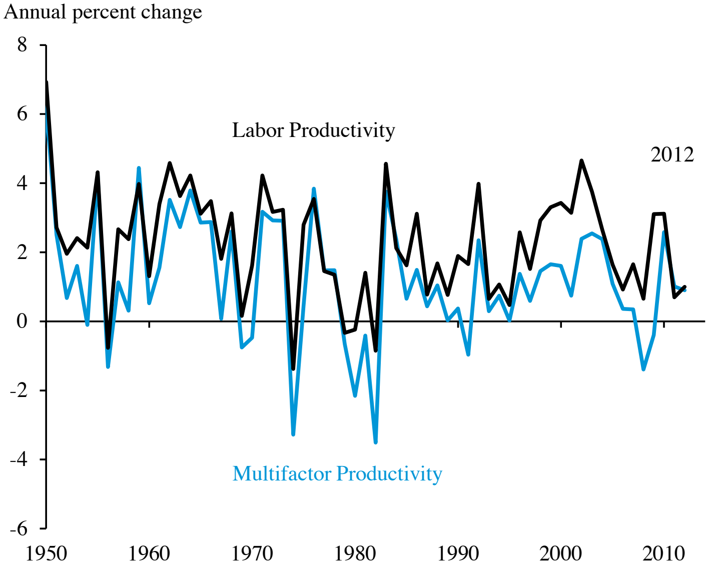
contribution of changes in labor quality (as measured by changes in the composition of the workforce), and the contribution of the growth in the amount of capital per worker.[4] The final column of Table 5-1 gives this decomposition, showing that 10 percent of the growth in labor productivity is due to improvements in the composition of labor (primarily greater educational attainment), 38 percent is due to increases in the amount of capital the worker has at his or her disposal, and 52 percent is due to increases in broad technological progress as measured by MFP.

The growth rates of the BLS measures of labor productivity and multifactor productivity have varied over time and are shown in Figure 5-1. Over the past 60 years, labor productivity has on average grown just over 1 percentage point faster than MFP: from 1953-2012, labor productivity grew at an annual rate of 2.2 percent per year, and MFP grew at an annual average rate of 1.1 percent per year.

As can be seen in Figure 5-1, both labor productivity and MFP are quite volatile from year to year. One reason for this volatility is measurement error in the estimation of both series; indeed, proper measurement of the inputs and outputs is a daunting task and for this reason alone not too much should be read into the growth of productivity in any one year. Another reason is that these series, and the gap between them, varies cyclically. For example, MFP growth fell—in fact, took on negative values—during the recessions that started in 1969, 1980-81, 1990, and 2007. These negative values do not mean that, during recessions, firms make negative technological

[4] Suppose aggregate production can be represented by the Cobb-Douglas production function, $Y = AL^{\alpha}K^{1-\alpha}$, where Y is real output, L is labor input measured in labor-quality units, K is capital, and A summarizes the contribution of technology to production, that is, A is TFP, and α is a constant. Then output per worker-hour (H) is $Y/H = A(L/H)^{\alpha}(K/H)^{1-\alpha}$. Thus the annual growth of output per worker, that is, the growth of labor productivity, is the sum of the growth of A, that is, the growth of TFP, plus α times the growth of L/H, that is the growth of labor quality per worker-hour, plus 1-α times the growth of K/H, that is, the growth of the capital-labor ratio. By using Tornqvist aggregation, the BLS MFP measure allows shares (α) to change over time and does not require an aggregate Cobb-Douglas production function.

Figure 5-1
Nonfarm Private Business Productivity Growth, 1949 –2012



Source: Bureau of Labor Statistics, Productivity and Costs, Multifactor Productivity.

progress or collectively forget about the innovations they have produced over the preceding years. Rather, such declines in MFP could come about from changes in relative prices, so that existing methods of production are no longer the optimal way to combine inputs to produce output. Negative MFP growth can also arise from variation in the utilization rates of capital and labor over the business cycle.

From the perspective of policies to foster long-term economic growth, these annual and cyclical fluctuations are less relevant than long-term trends in the growth rates of productivity. Figure 5-2 shows a centered 15-year moving average of the growth rates of labor productivity and MFP; and, Table 5-2 summarizes the compound annual growth rates of these series over 10- and 20-year periods ending in 2012, as well as the 60-year period from 1953-2012.

Table 5-2 and Figure 5-2 tell a similar story, which has two parts. First, over the long run the gap between labor productivity growth and MFP growth has fluctuated in a small range, with a difference of between 1.0 and 1.3 percentage points in decadal averages. Moreover, there is no noticeable trend in this gap: the mean difference in the growth rates of these two productivity measures over 2003-12 is within 0.2 percentage point of the mean difference over 1953-62. The stability of the difference between these

Table 5–2
Nonfarm Private Business Productivity Growth

Period	Multifactor Productivity	Labor Productivity	Difference
10–year Average annual rates of change			
1953–1962	1.5	2.6	1.1
1963–1972	1.9	2.8	1.0
1973–1982	–0.1	1.1	1.2
1983–1992	1.1	2.2	1.1
1993–2002	1.1	2.4	1.3
2003–2012	0.9	1.9	1.0
20–year Average annual rates of change			
Period	Multifactor Productivity	Labor Productivity	Difference
1953–1972	1.7	2.7	1.0
1973–1992	0.5	1.6	1.1
1993–2012	1.0	2.1	1.1
60–year Average annual rates of change			
1953–2012	1.1	2.2	1.1

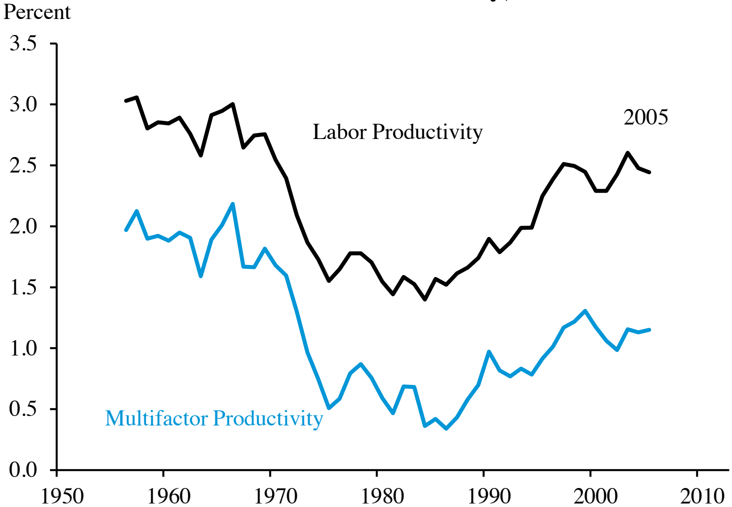
Source: Bureau of Labor Statistics, Productivity and Costs, Multifactor Productivity.

measures underscores the role of broad technological change—as measured by MFP—as a key driver of long-term growth of output per worker.

Second, over the past 60 years the long-term mean growth rates of labor productivity and MFP have varied substantially, in what appear to be three episodes. The first episode, the 1950s through early 1970s, experienced high growth of MFP (and of labor productivity), with MFP growth averaging 1.7 percent per year from 1953 through 1972. The second episode, the late 1970s through early 1990s, experienced much lower MFP growth, averaging 0.5 percent per year. The third episode, from the mid-1990s through the present, experienced an intermediate level of MFP growth of 1.0 percent per year.

Because productivity is the key to raising output per person, a great deal of academic research has focused on understanding why productivity growth varies over time. Research points to several factors that contributed to the productivity slowdown of the 1970s. A major culprit seems to be the sharp rise in energy prices during the 1970s that made less energy-intensive technologies more attractive, thus changing the optimal way to combine inputs and reducing MFP growth (Jorgenson 1988, Nordhaus 2004). One lesson learned from this period is how important energy cost fluctuations are in determining the growth of potential output.

Figure 5-2
15-Year Centered Moving Average of Annual Growth Rates for Labor and Multifactor Productivity, 1956–2005



Source: Bureau of Labor Statistics, Productivity and Costs, Multifactor Productivity; CEA Calculations.

Another explanation is due to rapid changes in the labor force in the 1970s, primarily shifting the workforce to newer, less-experienced workers. The Baby Boom generation (the cohort born between 1946 and 1964) came of age in the 1970s and 1980s, lowering the overall work experience of the economy. This was a period of rapid entry of women into the workforce for the first time, a shift that also temporarily reduced the overall level of workforce experience in the economy (Feyrer 2007, 2011). Moreover, the rapid entry of these new workers into the workforce outpaced investment, slowing the growth of the capital-labor ratio.

Another possible part of the story is that productivity growth in the 1950s and 1960s was temporarily spurred by large public investments such as the interstate highway system and the commercialization of military innovations from World War II like the jet engine and synthetic rubber.

The productivity rebound of the 1990s and 2000s is widely attributable to the information technology (IT) revolution. For the nine years from 1996 to 2005, MFP grew at 1.6 percent per year, a rate not seen in a nine-year period since the mid-1960s. Although many of the basic technologies that facilitated this growth, like the personal computer and the software to run it, were invented in the 1970s and 1980s; improvements in speed, breadth of applications, and the ability of firms to exploit this technology stretched through the ensuing decades. The BLS MFP measure suggests that much of the productivity improvement resulted from technological and process improvements, a position supported, for example, by Basu, Fernald, Oulton, and Srinivasan (2004). Alternatively, Jorgenson (2001) and Jorgenson and Ho (2012) emphasize the importance of the accumulation of physical IT capital. Oliner, Sichel, and Stiroh (2007) provide a detailed review of the literature on the 1990s productivity boom.

A key current question is what the rate of productivity growth will be going forward—will the U.S. economy maintain the pace of recent decades, will new innovations accelerate the pace of productivity growth, or will productivity growth revert to the slower rates before the recent boom? MFP growth fell sharply in the recession, grew sharply in the early stages of the recovery, and has averaged 1 percent for 2011 and 2012. These large cyclical swings make it difficult to assess whether there has been a recent change in the rate of technological progress, relative to the late 1990s and early 2000s. The academic literature reaches mixed findings concerning whether the IT productivity boom was temporary.[5] This literature also requires qualification because it predates the substantial data revisions to historical GDP and productivity that were released in the summer of 2013, which substantially revised upwards estimated productivity growth in some years in the 2000s.

Some contributions to this debate look further into the future. While some economists predict labor productivity growth could decline in coming decades because the scope for future transformative general-purpose inventions is limited (Gordon 2012), others argue that IT is in fact a general-purpose invention and, at least in the medium run, presents an ongoing stream of opportunities for workplace reorganization and efficiency gains, as well as spin-off technologies and improvements.[6] Bernanke (2012) argued that making these improvements often requires more than just purchasing hardware and software, and realizing potential productivity gains can require changes within and between organizations and thus take a considerable time to be fully realized.[7]

Ultimately, it is very hard to predict future growth rates in innovation, and there is no economic reason that these growth rates should be constant over time. Moreover, the past four decades have seen substantial changes in

[5] The findings in the literature on recent productivity growth trends tend to depend on the statistical approaches used to discern different productivity regimes. Authors that adopt discrete breaks or regime shifts, including Kahn and Rich (2011) and Fernald (2012), tend to conclude that the productivity growth boom has passed, whereas Oliner, Sichel, and Stiroh (2007), who use methods in which productivity growth evolves more slowly, find less of a slowdown.

[6] An example of such now-possible workplace reorganization is telecommuting; see, for example, Bloom, Liang, Roberts and Ying (2013); Noonan and Glass (2012), Bailey and Kurland (2002); and Busch, Nash, and Bell (2011).

[7] These issues are argued in the February 2013 TED debate between Gordon and Brynjolfsson.

the extent to which productivity gains translate into higher incomes across the board, the topic of the next section.

Productivity Growth and Inequality Growth

Productivity improvements provide more output that has the potential to benefit society broadly. Through the early 1970s, productivity gains led to increases in labor compensation. Since then, however, productivity growth has not translated into commensurate growth in labor compensation, and income inequality has increased markedly.

Trends in Inequality, Productivity Growth, and Compensation

Real output per hour was 99 percent higher by the end of 1972 than in 1947, while real average hourly earnings (GDP deflator) grew by 73 percent. Figure 5-3 shows that since the early 1970s, the paths of labor productivity and average hourly earnings diverged more widely. As a result, by the end of September 2013 real output per hour was 107 percent higher than at the end of 1972, but average hourly earnings had only grown 31 percent.[8]

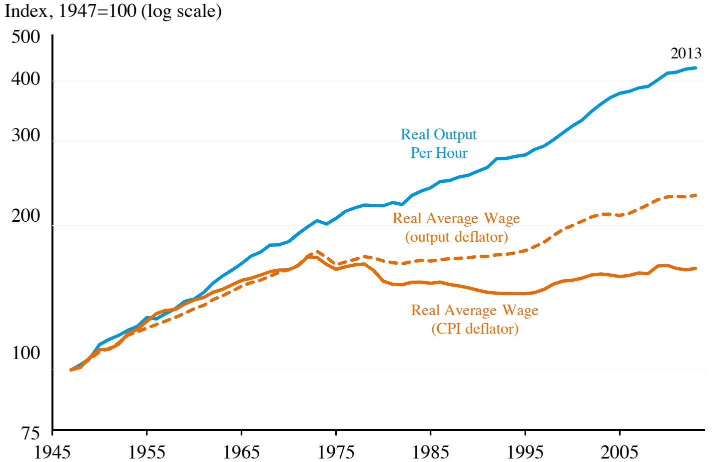
Table 5-3 examines the real output per hour and average hourly earnings for private production and nonsupervisory workers by decade. From 1953 to 1962, productivity growth exceeded the average annual rate of change in hourly earnings by only 0.4 percentage point. In the next decade, the difference in growth had ticked up to 0.6 percentage point. However, from 1973 through 2012 labor productivity grew 1.4 percentage points faster than earnings.

Since the 1970s, these trends generally have been worse for lower-income households than for higher-income households (DiNardo, Fortin, Lemieux 1996; Piketty and Saez 2003; Lemieux 2008; CEA 2012; Haskel, Lawrence, Leamer, and Slaughter 2012).[9] In particular, the income growth in the top percentile of the income distribution has been much stronger than other percentiles. For example, the Congressional Budget Office (CBO 2011) reports that from 1979 to 2007, real before-tax income at the median of the household income distribution increased by about 19 percent, while

[8] An alternative series from BLS measures real total hourly compensation (CPI deflator) for all nonfarm workers. This measure includes benefits as well as earnings. Since 1972, total hourly compensation has increased more than hourly earnings, but still only by 46 percent. BLS decompositions of compensation into real wage and benefit shares have been available since 1991. Since then, real wages grew 7 percent and benefits grew 22 percent, with the strongest benefit growth in the magnitude of employer contribution to health insurance.

[9] Figure 6-2 of this report suggests that the relative slow growth in income of the lower quintiles may have subsided some recently, particularly during the Great Recession and its near-term recovery. It is too soon to tell whether this has any implication for longer-term trends.

Figure 5-3
Growth in Productivity and Average Wage, 1947–2013



Note: Real output per hour is for all workers in the nonfarm business sector. Average wage is for private production and nonsupervisory workers. Output deflator is the price index for nonfarm business output. CPI deflator is the CPI-W. Data on wages before 1964 reflects SIC-based industry classifications.
Source: Bureau of Labor Statistics, Productivity and Costs, Current Employment Statistics; CEA calculations.

incomes for the top 1 percent of households have increased by around 200 percent.[10]

Technological Change and Inequality

The lesson from Figure 5-3 is that productivity growth is important for wage growth, but that does not mean that it automatically leads to wage growth. One possibility is that the sources of labor productivity and MFP growth since the early 1970s are qualitatively different than earlier, and that these different sources of growth drove the trends in inequality over the last 40 years. In the early 1990s, a broad consensus emerged among economists that an increase in the demand for skill relative to the supply of educated labor was the primary driver of the sharp rise in inequality in the 1980s (Bound and Johnson 1995; Katz and Murphy 1992; and Juhn, Murphy, and Pierce 1993). It soon became accepted that "skill-biased technological change" (SBTC) was the most important cause of increased inequality (Berman, Bound, and Griliches 1994; Krueger 1993). The crux of the argument is that, as computer technology became increasingly less expensive, relative demand increased for workers with complementary skills. This explanation has remained popular among economists with few modifications to the basic argument until recently (for example, Acemoglu 2002).

[10] The CBO notes that it chose 1979 and 2007 as points of comparison because there are cyclical fluctuations in inequality measures and both years are business cycle peaks.

Table 5–3
Average Annual Rates of Change in the Nonfarm Business Sector

Period	Real Output per Hour of all Workers	Average Hourly Earnings for Private Production and Nonsupervisory Workers	Difference (p.p.)
1953–1962	2.5	2.1	0.4
1963–1972	2.7	2.1	0.6
1973–1982	1.1	−0.4	1.4
1983–1992	2.2	0.4	1.8
1993–2002	2.3	1.8	0.5
2003–2012	2.1	1	1.1

Note: Both series are deflated by the price index for output in the nonfarm business sector. Data on earnings before 1964 reflect SIC–based industry classifications.
Sources: Bureau of Labor Statistics, Productivity and Costs, Current Employment Statistics; CEA Calculations.

While this hypothesis has remained influential, there are reasons to question the primary role of technology in causing the inequality changes that emerged in the 1980s. For example, many other industrialized nations, such as Germany and Japan, experienced similar technology shocks in the 1980s, but saw little or no increase in wage inequality. This led some economists to expand the framework for explaining inequality to acknowledge the importance of wage-setting institutions in mediating technology shocks (Freeman and Katz 1995). This critique gained more force with other researchers finding that changes in institutions—especially the decline in the real value of the minimum wage and labor unions—could account for much of the rise in inequality in the 1980s, at least in the bottom of the distribution (Lee 1999 and DiNardo, Fortin, and Lemieux 1996). An additional challenge to the skill-biased technological change hypothesis is that the timing of changes in inequality do not line up well with the nature of technological change across decades. Inequality in the bottom of the distribution rose in the 1980s, but has been flat or declining since then. However, much of the widespread business adoption of IT, including the Internet, occurred in the 1990s, and those innovations were at least as significant as the changes in the 1980s (Card and DiNardo 2002). In fact, inequality in the top of the distribution did continue to rise, but after rising sharply in the 1980s, inequality at the bottom of the distribution has been flat or declining since.

Goldin and Katz (2008) focus on changes in the growth of the supply of skills rather than on episodic increases in technological change. Using the ratio of college to non-college workers as a measure of the relative supply of skills, they show this relative skill supply grew by 3.9 percent from 1960 to 1980. But in 1980, as confirmed by Heckman and LaFontaine (2010) and others, this increase slowed as high school graduation rates stopped improving and college completion rates slowed. Goldin and Katz (2008) show that a constant increase in the demand for relative skill, combined with the post-1980 slowdown in the supply of relative skill, explains the time path of the logarithm of the college wage premium, which is one measured aspect of wage inequality.[11] The nature of rising wage inequality started to change around the early 1990s, becoming increasingly concentrated in the top end of the wage distribution. The ratio of the 90th to the 50th percentile of the wage distribution continued to grow at roughly the same rate it had since the early 1980s, whereas inequality at the bottom (the 50-10 ratio) declined somewhat after the late 1980s. Piketty and Saez (2006) find that income gains have increasingly concentrated in the top 10 percent and top 1 percent since the 1980s. The result has been a "polarization" or a "hollowing out" of the wage distribution, with relative wage growth in the bottom and especially the top of the wage distribution relative to the middle (Goos and Manning 2007; Autor 2010; Acemoglu and Autor 2011; Lemieux 2006).

Autor and coauthors refine the earlier skill-biased technological change literature and argue that the changes in inequality are driven by technological change that substitutes for some tasks but not others (Autor, Levy, and Murnane 2003; Autor, Katz, and Kearney 2006; Acemoglu and Autor 2011). In particular, this new research argues that computer technologies complement non-routine cognitive tasks, which tend to be highly paid; substitute for routine tasks, which tend to be in occupations with wages in the middle of the distribution; and have little effect on manual tasks that tend to be associated with lower wages. This technological explanation for polarization has been controversial, however, and Mishel, Shierholz, and Schmitt (2013) suggest that the theory does not explain the timing of changes in polarization, and more generally that occupational employment and wage trends do not explain a large part of the trends in wages or inequality over time. Moreover, one of the most striking changes in inequality over the past three decades—the sharp growth of incomes at the very top of the distribution—is unlikely to be related to technological changes or to a relative demand for skill (Alvaredo, Atkinson, Piketty, and Saez 2013).

This discussion has focused on whether increases in productivity translate into increases in earnings or lead to increasing inequality. A related, less-understood question is whether increasing inequality might

[11] This theory is based on evidence from before 2008. The U.S. economy has long had some skills shortage, which tended to turn up in the form of wage differentials rather than unemployment. It does not account for the large shock in aggregate demand that characterized the Great Recession, or the shock-driven unemployment rates from which the economy is still recovering.

directly dampen productivity growth, and this question is addressed further in Box 5-2.

Policies to Foster Productivity Growth and to Help Ensure That Everyone Benefits from it

The benefits of technological progress do not accrue only to those who develop new processes and inventions; they also spill over to the population at large. For this reason, the U.S. Government has a role in supporting and enabling technological development. This government role includes: directly funding or providing incentives for research and development (R&D); providing an institutional, legal, and regulatory environment that protects competition, defines and supports intellectual property rights, and thereby encourages private innovation; and developing human capital through education, especially in scientific and technological fields. In addition, the government has a role in ensuring that everyone benefits from those technological advances.

Investments in R&D often have "spillover" effects; that is, a part of the returns to the investment accrue to parties other than the investor. As a result, investments that are worth making for society at large might not be profitable for any one firm, leaving aggregate R&D investment below the socially optimal level (for example, Nelson 1959). This tendency toward underinvestment creates a role for research that is performed or funded by the government as well as by nonprofit organizations such as universities.

These positive spillovers can be particularly large for basic scientific research. Discoveries made through basic research are often of great social value because of their broad applicability, but are of little value to any individual private firm, which would likely have few, if any, profitable applications for them. The empirical analyses of Jones and Williams (1998) and Bloom et al. (2012) suggest that the optimal level of R&D investment is two to four times the actual level. Akcigit et al. (2013) also find underinvestment in basic research (although, contrary to the bulk of the literature, they find overinvestment in applied research), and suggest policies that are specifically targeted at basic research.

Consistent with the presence of large spillover benefits, most basic research in the United States is funded by the government and other nonprofit entities. As Figure 5-4 shows, over half comes from government sources, and less than one-quarter comes from private industry. However, expenditures on basic research are only a fraction of total R&D expenditures, as seen in Figure 5-5, and the private-sector share of funding for applied research and development is much higher than it is for basic research.

Box 5-2: Does Inequality Affect Productivity?

Although conventional economic models do not include the equality of the income distribution as a determinant of economic output, some recent research has focused on whether increasing income inequality might reduce the growth rate of productivity. There are at least three channels that could produce this link and, in each, an underlying source of income inequality potentially leads to slower productivity. The first channel is through disparities in access to, and the quality of, publicly funded secondary education: inequality in educational quality leads to disparities in skills, so an increase in labor hours might not increase labor quality, slowing labor productivity growth. For example, Goldin and Katz (2008) argue that in the 19th and early 20th centuries, greater access to education in the United States than in Europe resulted in the United States having higher rates of labor productivity growth. In the United States today, the relevant channel is not likely related to access to public schools, but more likely geographic disparity in resources available to students at those schools.

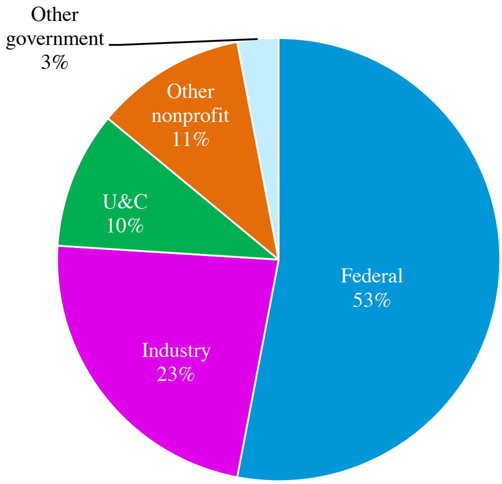
A second channel is that greater income inequality creates disparities in the ability to pay for privately funded education, especially pre-kindergarten and college.[1] This channel too is relevant because of the increasing expense of post-secondary education.

A third channel, discussed by Acemoglu and Robinson (2011), is that sufficiently powerful and entrenched elites have an incentive to use resources to protect their interest rather than encourage growth. The relevance of Acemoglu and Robinson's examples of extractive societies drawn from world history—ancient Rome, the Mayans, slave-dependent economies in the early Americas, and so forth—to the United States today is less clear than that of the other channels.

There have been some attempts to use cross-country differences as sources of variation for econometric studies of the link from inequality to productivity growth. Those attempts, however, confront a variety of data availability and measurement issues, including comparable measures of inequality (Fields 2001) and insufficient variables to avoid spurious effects being loaded onto the inequality measure (Banerjee and Duflo 2003). In any event, the question of whether the increases in U.S. inequality over the past two decades have dampened, or could dampen, productivity growth remains an important source of concern.

[1] Except for programs like Head Start, pre-kindergarten education is privately financed. Heckman, Pinto and Savelyev (2012) contribute and list literature demonstrating the importance of early childhood intervention to subsequent schooling and other life outcomes. At the college level, nearly all students pay at least some of their educational expenses.

Figure 5-4
Basic Research Expenditures in the U.S. by Source Funding, 2010



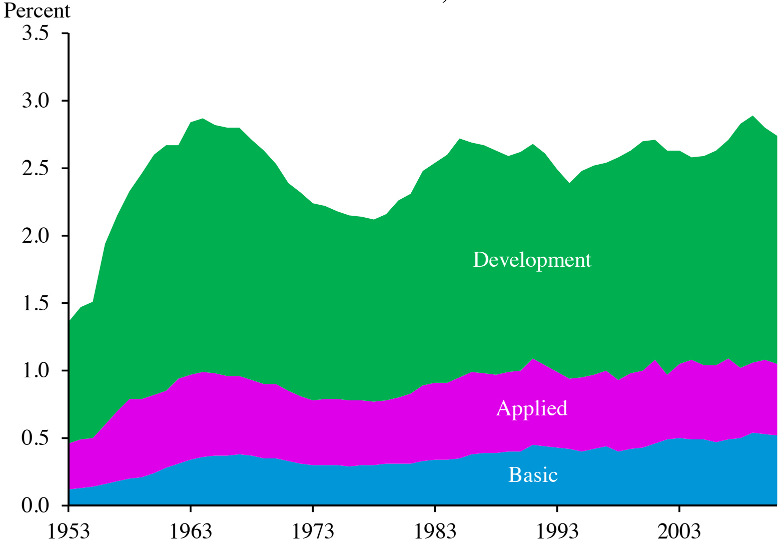
Source: National Science Foundation, National Patterns of R&D Resources: 2010-11 update.

In addition to direct funding of R&D, the government also provides financial incentives for private R&D investment through tax policy. Government can also facilitate private R&D investment and technological progress by providing an institutional, legal, and regulatory framework that clarifies and enforces intellectual property rights and thereby ensures that innovators reap enough financial rewards from their innovations to provide sufficient incentive to engage in a closer-to-optimal level of R&D.[12]

One important type of intellectual property right is patents. A patent grants the inventor a temporary exclusive right over the invention. Exercising that right results in high prices and profits for investments that are successfully commercialized, and those profits provide an incentive to invent. However, the exercise of the exclusive right will also raise prices on inventions that would have been created even with weaker patent protection or with none at all, and these higher prices harm consumers. Moreover, because patented inventions are sometimes used as inputs in creating additional innovations, the higher prices created by patents (as well as the associated legal and administrative burdens, such as negotiating licenses) could slow down subsequent innovation. As discussed further below, a central economic challenge of patent policy is to strike the right balance

[12] Research in development economics suggests that a key factor in the economic performance of a country is its "institutions," such as rule of law and clear property rights (Hall and Jones 1999, Rodrik et al., 2004).

Figure 5-5
Composition of Total R&D Spending as a Share of GDP, 1953–2011



Source: National Science Foundation, National Patterns of R&D Resources: 2010-11 update.

between providing an economic incentive to invent and the potential harm from the exercise of patent rights. At a minimum, it is important to ensure that patents are not wrongly issued, but rather are only issued for inventions that are non-obvious, useful, and inventive.

The government can also lay the groundwork for greater creativity and invention by supporting the development of human capital. Investments and improvements in education and training, particularly in the Science, Technology, Engineering, and Mathematics (STEM) fields, foster the innovation workforce of the future.[13] The productivity of these workers can be enhanced by investment in "innovation clusters," which are dense concentrations of firms and of highly skilled personnel, usually close to a major research university, whose mutual proximity can further promote innovation (see Greenstone, Hornbeck, and Moretti 2008).

Immigration reform is another human capital policy that has the potential to increase the pace of innovation. Studies have found that foreign-nationals living in the United States authored or co-authored over 25 percent of U.S. patent applications in 2006, and that over 75 percent of patents awarded to the top 10 patent-producing American universities in 2011 had at

[13] As discussed in Delgado et al. (2012), one determinant of a country's economic performance is its science and innovation infrastructure. The authors include in this category a number of elements that can be influenced by supportive government policy, such as the quality of scientific research institutions and the quality of math and science education.

least one foreign-born inventor. Moreover, the innovation benefits of immigration are not confined to the immigration of innovators. Immigration of low-skilled workers, as well as immigration of high-skilled workers who are not innovators, can spur innovation indirectly by increasing specialization. When more non-innovators are present to specialize more completely in their occupations, they enable innovators to specialize more completely in theirs. The Congressional Budget Office (2013) projects that the additional immigration resulting from the Border Security, Economic Opportunity, and Immigration Modernization Act, as passed by the Senate, would raise total factor productivity by roughly 0.7 percent in 2023 and by roughly 1.0 percent in 2033 as a result of increased innovation and task specialization.

Finally, the government has an important role in ensuring that access to the technologies that catalyze productivity growth, and to the technologies and products that are the fruits of that productivity growth, are broadly available throughout American society. Sharing these benefits increases welfare directly, and also ensures that the broad population maintains the technological skills needed in the workplace and for the education of current and future generations.

This chapter now turns from a general discussion of the role of government policy in achieving technological progress to a focus on two key current areas that are important for productivity growth and that are also a focus of the Administration's policies: telecommunications and patent reform.

Telecommunications and Productivity Growth

The telecommunications industry is an important one for fostering productivity growth. Improved telecommunications infrastructure, particularly fast and widely accessible wired and wireless broadband networks, is a critical factor in enabling important technological advances in business, health care, education, public safety, entertainment, and more. Government policies have an important role to play in facilitating and catalyzing these improvements, as discussed below. In this chapter, telecommunications policy is discussed in particular detail, in part due to its importance, and in part because it serves as a good illustration of more general economic and policy principles.

Innovation and Investment

The telecommunications sector is a major success story in the U.S. economy. A recent White House (2013) report, *Four Years of Broadband Growth*, documents many of the striking facts, including:

- Just two of the largest U.S. telecommunications companies account for greater combined stateside investment than the top five oil/gas companies, and nearly four times more than the big three auto companies combined, as seen in Figure 5-6.
- Between 2009 and 2012, annual investment in U.S. wireless networks grew more than 40 percent, from $21 billion to $30 billion. During that period, investment in European wireless networks remained flat, and wireless investment in Asia (including China) rose only 4 percent. The report projected that U.S. wireless network investment would increase further in 2013, to $35 billion.
- The United States leads the world in the availability of advanced 4G wireless broadband Internet services such as LTE; nearly half of the global subscriber base for 4G LTE is in the United States.
- The United States ranks among the top countries in the world in the amount of currently licensed spectrum available for mobile broadband.

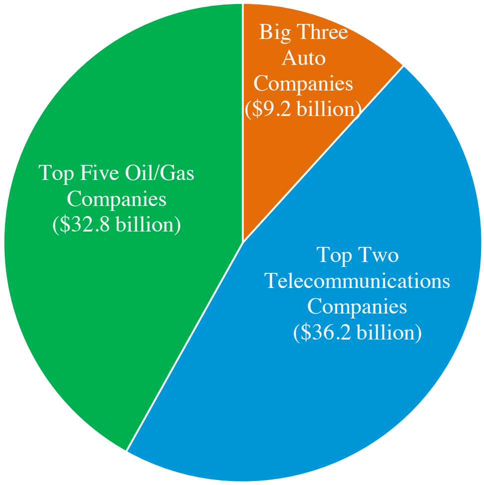
This infrastructure is at the center of a vibrant ecosystem that includes smartphone design, mobile applications development, and the use of these technologies to effect broader changes in the economy and society—all of it centered in the United States. The mobile applications industry is forecast to generate more than $25 billion in revenue in 2013, rising to $74 billion in 2017, with nearly 2 million applications available for download at the two largest mobile app stores. Improved telecommunications has also contributed to changes in the way that business is organized, and in ways that may lead to further improvements in productivity. An example of this is discussed in Box 5-3.

Four Key Areas for Telecommunications Policy

The U.S. Government can support innovation and investment in telecommunications through the same general policies discussed above: direct government investment in research and development; catalyzing private innovation through policies such as reforming and extending the Research and Experimentation Tax Credit; catalyzing technological infrastructure investment in areas like broadband; and ensuring that everyone benefits from broadband technologies.

Government Investments in Research and Development. As discussed above, spillover benefits to research and development, especially for basic science and technology, creates a role for direct government investment. Perhaps the most famous government investment in telecommunications technology was the Defense Advanced Research Projects Agency (DARPA) development of the Internet. But DARPA has provided other important defense-based public research contributions as well. These contributions

Figure 5-6
Relative Investment of the Telecommunications Sector, 2011



Source: Progressive Policy Institute.

include the radio and, more recently, Global Positioning Systems, which today are central to a huge number of consumer applications.

Today, the Department of Defense (DOD) continues to play an important role in telecommunications research, particularly in helping to develop ideas and technologies for sharing of electromagnetic spectrum frequency bands between different users, including between government and private users. This, as discussed further below, has been identified as important for efficient spectrum management in the future (PCAST 2012). For example, DOD has solicited innovative research proposals aimed at efficient and reliable sharing of spectrum between radar and communications systems. All told, $100 million in Federal investments are being targeted toward spectrum sharing and advanced communications through the National Science Foundation (NSF), DARPA, and the Commerce Department.

Catalyzing Private Investment. Reforming, expanding, and making permanent the Research and Experimentation Tax Credit would increase investment in telecommunications technology, accelerating innovation. Immigration reform would accelerate innovation as well. Reforming the patent system is also important in this industry, especially for technology deployed in smartphones, which are complex devices that embody thousands of patents. The increasing frequency of patent disputes in this area suggests that there may be increasing costs to navigating the appropriate

Box 5-3: Just-in-Time Manufacturing

The just-in-time (JIT) approach to manufacturing aims to maximize profits by dramatically reducing inventories and their costs. By minimizing the time that inventory is held, the system allows a fixed amount of inventory space to be used more productively; that is, by processing more goods through the fixed spaced during a fixed amount of time. Agrawal (2010) delineates channels through which the JIT approach can reduce costs, improve quality and customer service, maintain flexibility, and promote logistical efficiency. Many of these channels now rely on improved information and telecommunications technology. Since JIT requires precise coordination between demand and supply, the contemporaneous tracking of each is essential. On the supply side, Zhang et al. (2012) argue that radio frequency identification technology can provide firms with precise, accurate, real-time information on materials as they pass through the manufacturing process. But technology that makes JIT feasible is only one requirement. Other studies (Hur, Jeong and Suh 2009, Tayal 2012, Fairris and Brenner 2001, Agrawal 2010, Sim and Koh 2003) show that organizational experimentation, innovation, and learning in using the technology can also be necessary to realizing productivity gains.

licenses. If these costs are high enough to adversely affect the introduction of new products, then patent reform is particularly important for the telecommunications industry.

Catalyzing Technological Infrastructure Investment. The Federal Government funded the country's first investment in telecommunications infrastructure, a telegraph line from Washington D.C. to Baltimore built in the 1840s. But since then, appropriately, the vast majority of technological infrastructure investment has been private. Over the course of decades, an extraordinary expansion of telecommunications infrastructure made basic telephone service available to nearly every resident of the country, far sooner than in most other countries, which is a remarkable achievement given the large size and relatively low population density of the United States.

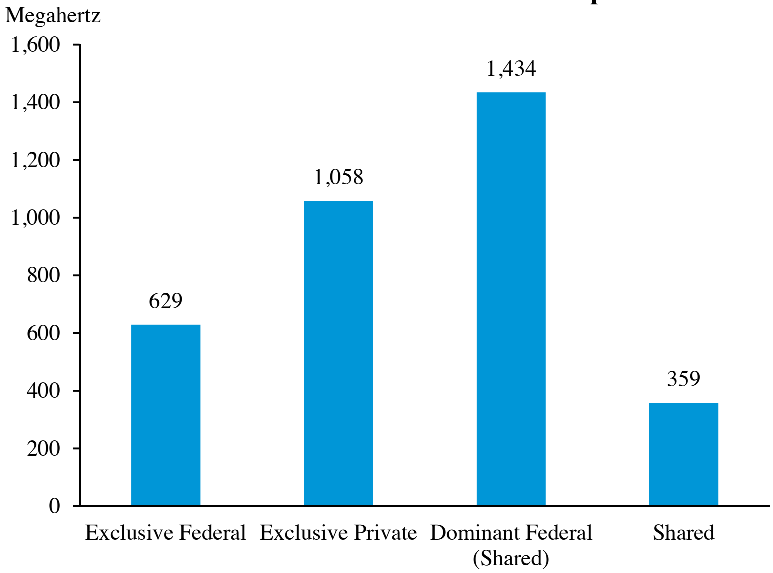
Public policy encouraged these investments. Many private carriers, as regulated monopolies, were permitted to charge high rates for long-distance calls, business service, and the telephones themselves. A portion of the resulting funds were required to be used to subsidize basic local phone service, particularly in rural and other areas that are costly to serve due to low population density and geographic factors. The Telecommunications Act of 1996 sought to reform and improve upon telecommunications regulation by enabling greater competition, particularly in local and long-distance telephone service, and by rationalizing, and making explicit, the subsidy system supporting service in high-cost areas. Substantial additional private investment followed.

In recent years, the U.S. Government has further facilitated private telecommunications investment through favorable tax policy. In 2010, the President proposed and signed into law the largest temporary investment incentive in history—100-percent expensing—that, together with the bonus depreciation that preceded and followed it, played a critical role in increasing and accelerating investment, including the substantial increases in both wired and wireless investment in the telecommunications sector. For example, two major companies in a joint statement said that, "despite the downturn in the economy, the cable communications sector has been able to continue steady investment and to retain jobs as a result of policies like 100-percent expensing."

Catalyzing investment in mobile broadband infrastructure is especially important given the rapidly growing usage and the scarcity of electromagnetic spectrum for carrying wireless broadband traffic. In 2010, Federal Communication Commission experts predicted that the needs for broadband capacity will overwhelm available spectrum (the "spectrum crunch"). If allowed to happen, this would result in higher prices for mobile broadband services, as well as reduced growth in broadband-based innovation, services, and employment. The scarcity of broadband capacity can be alleviated through increased investment (denser transmission infrastructure means more traffic on a given spectrum frequency), fuller deployment of spectrum already licensed to wireless carriers, spectrum license consolidation, technological advancement, and improvements in spectrum policy.

One important initiative is to seek to reallocate public spectrum when it has a more valuable private use. The Federal Government is a major user of spectrum, as Figure 5-7 shows. Most of this usage involves national security and law enforcement functions, as shown in Figure 5-8. Federal use of spectrum is valuable, but it is not costless. As an economic matter, if a particular spectrum band would produce a larger net social surplus in private hands than in public hands, then it should be reallocated, and vice-versa. That is, the Federal Government can alleviate spectrum scarcity by having government agencies vacate certain spectrum bands entirely, or share them with private users, when this can be achieved without compromising the agencies' vital missions (which in many cases involve safety-of-life and national security) and when the associated costs of relocating government operations out of those bands are justified by the social value that will be unlocked as a result of the reallocation to the private sector. The vacated

Figure 5-7
Exclusive and Shared Allocation of Radio Spectrum



Source: National Telecommunications and Information Administration (2009).

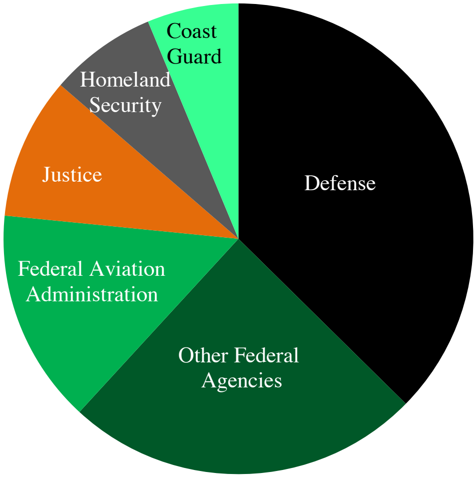
spectrum could then be auctioned off to commercial users or, if appropriate, made widely available on an unlicensed basis (more on this below). The Federal user would be relocated to alternative spectrum that could be used more intensively and economically, particularly if additional resources were made available for investment in newer equipment.

In addition to economizing on its aggregate spectrum usage, the government can further alleviate spectrum scarcity by rationalizing spectrum allocation. There are some spectrum bands that, above and beyond the properties that make them valuable in general (for example, strong propagation through buildings and in rural areas), are particularly valuable for commercial applications, such as if they are complementary to other commercial spectrum bands. In those cases, value can be unlocked by having the government relocate from those bands to other bands that do not have that property—again, under the condition that this can be done without compromising vital missions and that the relocation costs are not prohibitively high.

Box 5-4 describes several spectrum investment policies that have been undertaken or proposed by the Administration.

There is also substantial scope to reallocate some spectrum currently licensed to private entities to a more valuable use in wireless broadband. Some incumbent firms, such as over-the-air broadcast television stations, hold rights to spectrum that are much more valuable as wireless broadband

Figure 5-8
Federal Agencies with Most Spectrum Assignments



Note: "Other Federal Agencies" includes Interior, Agriculture, Energy, Commerce, and the other remaining 48 agencies and departments with spectrum frequency assignments.
Source: National Telecommunications and Information Adminstration, Government Master file (2010); Government Accountability Office Analysis.

spectrum. The 2010 National Broadband Plan introduced the idea of "incentive auctions" as a tool to help meet the nation's spectrum needs by giving those rights-holders a share of the auction proceeds if they relinquish their rights. In the Spectrum Act of 2012, Congress authorized the Federal Communications Commission to conduct incentive auctions and directed that the FCC use this innovative tool for an incentive auction of broadcast television spectrum. In September 2012, the FCC adopted a Notice of Proposed Rulemaking in order to develop a rulemaking record that will enable the Commission to meet the challenges presented by the Spectrum Act's unique grant of authority. The magnitude of potential gains to social surplus are enormous when broadcasters with access to new, more-efficient transmission technologies that use less spectrum, or with a small and shrinking base of over-the-air viewers and annual revenue in the low millions of dollars, will have an incentive to relinquish spectrum that, when reconfigured for commercial broadband use, will be sold for hundreds of millions of dollars to companies that will use it to improve services for a vastly greater number of broadband customers.

Some spectrum can be used effectively without being licensed at all, but rather made available for anyone to use on an unlicensed basis. Just as some roads seldom experience traffic jams, in some instances certain spectrum bands do not become highly congested even when access is free.

Box 5-4: Spectrum Investment Policies

In 2010, the President issued a Presidential Memorandum called "Unleashing the Wireless Broadband Revolution," which directed the Secretary of Commerce, working through the National Telecommunications and Information Administration (NTIA), to collaborate with the Federal Communications Commission (FCC) to make available a total of 500 megahertz (MHz) of Federal and nonfederal spectrum over the next 10 years, suitable for both mobile and fixed wireless broadband use, nearly doubling the amount of spectrum available for such purposes. The Secretary of Commerce has been facilitating discussions between agencies and nonfederal entities that have produced an unprecedented level of information-sharing and collaboration to identify opportunities for agencies to relinquish or share spectrum, currently focusing on the 1695-1710 MHz band, the 1755-1850 MHz band, the 3550-3650 MHz band, and the 5350-5470 and 5850-5925 MHz bands.

The President's fiscal year 2015 budget would invest $7.5 million to monitor spectrum use by Federal agencies in high-priority markets to identify opportunities for repurposing spectrum through auctions, while protecting Federal missions. This budget proposal builds on the Middle Class Tax Relief and Job Creation Act of 2012 by proposing to authorize use of a spectrum license user fee for licenses not currently awarded via auctions (for example, international satellite licenses), to promote efficient utilization of spectrum. This fee would raise nearly $5 billion over the next 10 years, and would continue to encourage more efficient allocation and use of spectrum.

Unlicensed spectrum plays an important role in the broadband ecosystem, enabling Wi-Fi, Bluetooth, "smart homes," and more, which operate on unlicensed spectrum using devices whose power is low enough that interference among numerous devices sharing the spectrum is not a major concern. It also helps to alleviate scarcity in licensed spectrum bands. This is because a great deal of mobile usage is not the "on-the-go/in transit" mobile usage that must be transmitted on a carrier's licensed mobile network, but rather is so-called "nomadic" usage (for example, at home, office, or other fixed location), that is amenable to carriage mostly by a wired broadband connection and then wirelessly completed using a nearby unlicensed Wi-Fi router. For this reason, the licensed carriers are investing heavily in the deployment and use of Wi-Fi networks. The value of this unlicensed spectrum has been estimated at $16 billion to $37 billion per year.

In February, the FCC proposed to make available up to 195 megahertz of additional spectrum in the 5-gigahertz band for unlicensed wireless devices, a 35 percent increase. This band was selected for unlicensed use in part because the presence of incumbent users of this band, including automobile makers that have been developing short-range communications capabilities that could greatly improve traffic safety and efficiency, make it a poor candidate for being vacated and auctioned off for licensed use. To unlock the value of the band for unlicensed use, the FCC has also proposed to create a more flexible regulatory environment, and to streamline existing rules and equipment authorization procedures for devices throughout this band. Currently ongoing is the process of identifying regulatory changes that strike the best balance between unlocking the value of this spectrum for unlicensed use on the one hand, and avoiding harmful interference with incumbent users on the other.

Clearing Federal Government spectrum for exclusive licensed use, and making it available for shared unlicensed use, remain viable solutions in the near term. However, given the dramatic spectrum challenge and the fact that much of the lowest-hanging fruit for reallocation has already been picked, it is also important to focus on newer and more innovative ideas. These ideas include new advances in the sharing of spectrum between different users, particularly between government and private users. Innovation in spectrum sharing is both promising and necessary, as there are some spectrum bands that the government cannot vacate entirely, but that nevertheless have unused capacity, and that with appropriate processes and procedures in place could be shared, accommodating some valuable private usage without compromising mission-critical functions.

The President's Council of Advisors on Science and Technology (PCAST) released a report estimating that "in the best circumstances, the amount of effective capacity that can be obtained from a given band of spectrum can be increased thousands of times over current usage through dynamic sharing techniques that make optimal use of frequency, geography, time and certain other physical properties of the specific new radio systems (PCAST 2012)."

The 2010 Presidential Memorandum that set the Administration's spectrum goal contemplated the sharing of Federal Government spectrum as one means of achieving that goal. More recently, in June 2013, another Presidential Memorandum established a Spectrum Policy Team in the Executive Office of the President, which was charged with the mandate to "monitor and support advances in spectrum sharing policies and technologies." That Memorandum also contains measures to facilitate research, development, testing, and evaluation of technologies to enhance spectrum sharing and other spectrum-related efficiencies.

To stimulate investment in more advanced forms of spectrum sharing, the Defense Advanced Research Projects Agency (DARPA) is soliciting innovative research proposals aimed at efficient and reliable sharing of spectrum between radar and communications systems. Consistent with its history of promoting groundbreaking technological breakthroughs for both military and commercial use, DARPA is seeking "innovative approaches that enable revolutionary advances" in spectrum sharing, specifically in the spectrum bands that are most amenable to broadband and communications services. The program may fund multi-year projects designed either to significantly modify existing radar and communications systems or to unveil new system architectures redesigned from the ground up.

By itself, making additional Federal spectrum available for commercial use, whether on an exclusive or a shared basis, is unlikely to be sufficient to keep up with the exploding demand for bandwidth. The ambitious goal of freeing up 500 MHz of spectrum would nearly double the amount of wireless spectrum available for mobile broadband over the course of a decade, but even that may not be enough to keep up with spectrum usage growth. Therefore, it is important to do everything from increasing investments in wired broadband networks that can offload some of the demand (often making the last connection wireless, but through Wi-Fi rather than cellular), to increasing the density of wireless cells, to encouraging technological innovations for using spectrum more efficiently.

The Administration is trying to help with these efforts in a variety of other ways, including the June 2012 Executive Order issued by the President specifying a number of steps that will ease and facilitate carriers' access to Federal land and buildings for purposes of deploying broadband infrastructure, including cell towers.

Ensuring Everyone Benefits. It is important to ensure broad participation in the benefits of broadband telecommunications technologies, because broad participation allows more people to use those benefits to develop their talents, which lead to higher economic growth and higher living standards in the future. One element of broad participation is ensuring that technology and its products are affordable. To that end, vigorous antitrust enforcement is critical to ensure that that prices are not inflated and choices not limited by lack of competition. This has been a focus of the law enforcement agencies, and is also important as a policy consideration going forward.

The Obama Administration has made critical investments in expanding broadband to underserved communities. The American Recovery and Reinvestment Act of 2009 included $6.9 billion in funding to upgrade the nation's broadband infrastructure, with $4.4 billion administered by the Department of Commerce's National Telecommunications and Information

Administration, and $2.5 billion by the Department of Agriculture's Rural Utilities Service. Of these funds, a total of $4.4 billion (as of the end of May 2013) went to fund more than more than 325 broadband projects through the Broadband Technology Opportunities Program and the Broadband Initiatives Program. The Federal Communications Commission has also played an important role in expanding broadband deployment in unserved and underserved areas through Universal Service Reform and the establishment of a $4.5 billion annual Connect America Fund, which reallocates funds previously used to support voice service.

Education researchers have long believed that technology holds the potential to profoundly impact the classroom experience, from allowing students to interact with course content in new and personalized ways to helping teachers understand what lessons and techniques are most effective. By making the ever-expanding collection of educational resources available on the Internet accessible to teachers and students in classrooms, technologically equipped schools enhance learning by gaining access to those resources, rather than being limited to resources that are physically at hand.

Although more high quality research on the effectiveness of online educational tools is still needed, these tools do show promise. A meta-analysis of experimental or quasi-experimental studies of the effects of online education conducted by the Department of Education in 2010 found that students who receive instruction that combines online and face-to-face elements performed better than students who received either exclusively online, or exclusively face-to-face, instruction. Other factors such as instruction time or curriculum may contribute to this positive effect, but the meta-analysis suggests that further research on designing, implementing, and evaluating these blended approaches may be worthwhile.

Instruction methods that incorporate computers have also shown promise in mathematics education. Barrow, Markman, and Rouse (2009) found that students who were randomly assigned to participate in and complete computerized math lessons at their own pace scored 0.17 to 0.25 standard deviations higher on mathematics achievement tests than students who received traditional instruction. Computer-aided mathematics instruction has been shown to have similar effects in other contexts. In an experimental study, Banerjee et al. (2007) find that playing educational math games on computers for two hours a week improved the math scores of impoverished elementary school students in India by 0.47 standard deviations. In another experiment, Carillo, Onofa, and Ponce (2010) find poor Ecuadoran elementary school students who used adaptive math and language software for three hours a week improved their math scores by 0.30 standard deviations.

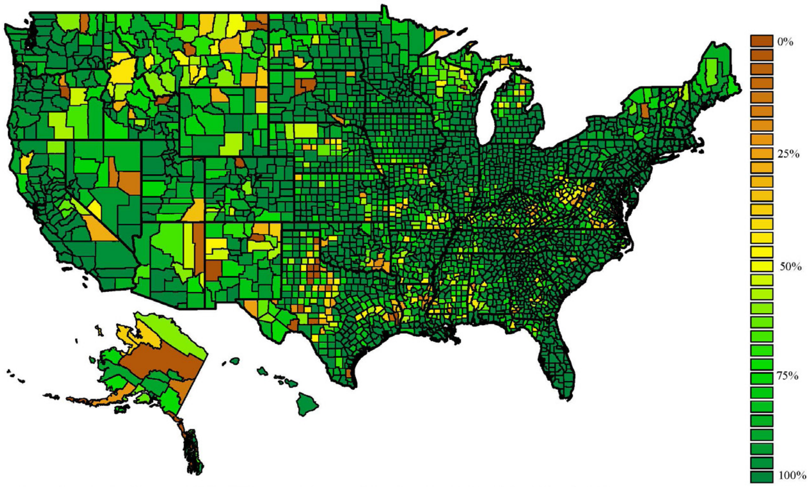
Because current uses of technology can enhance learning and the potential of future developments is untold, it is critically important that all students have access to 21st century classrooms. The ConnectED program, announced by President Obama in June 2013, takes important steps to ensure that the benefits of improvements in educational technology will be made widely available. While initiatives like the FCC's E-Rate program, established under the Telecommunications Act of 1996, have helped bring Internet access to almost every school in the nation, many schools do not have access to the fast broadband speeds enjoyed by most businesses and households. Further, the E-Rate program was designed primarily to bring Internet access to the school, with less priority on ensuring that access was available throughout the school, such as via Wi-Fi technology. As a result, 62 percent of school districts say their bandwidth needs will outstrip their connections within the next 12 months, and 99 percent say that this will happen within three years.

ConnectED will bring high-speed broadband and wireless Internet access to 99 percent of America's students in their school classrooms and libraries within five years. To make the most of this enhanced connectivity, ConnectED will refocus existing professional development funds to train teachers to take full advantage of these resources in order to improve student learning. Finally, by equipping schools with the broadband Internet access necessary to make use of high-tech educational devices, ConnectED will deepen the market for such devices, as well as the digital educational content with which they interact, spurring private-sector innovation in this area.

The President has called on the FCC to modernize the E-Rate program, and has also called on the expertise of the NTIA, in order to deliver this connectivity and meet the goal of connecting 99 percent of America's students to the digital age within five years through next-generation broadband and high-speed wireless in their schools and libraries. Answering that call, the FCC announced in February 2014 that it would invest $2 billion to connect 20 million students over the next two years, representing a crucial down-payment on reaching the President's goal. The initiative, however, is not just about infrastructure. The President announced in February over $750 million in private sector commitments to help fill out this vision of a connected classroom through the digital devices, content and learning software, home wireless access, and teacher training necessary to make the best possible use of this infrastructure. By leveraging all these resources, we are making substantial progress toward a world-class education for every student that does not depend on their family's income or on the zip code in which they were born.

Figure 5-9

Percentage of Households with Access to Download Speeds of 6 Megabit per Second or Greater



Source: Department of Commerce, National Telecommunications and Information Admnistration, National Broadband Map.

Finally, it is crucial that the benefits of broadband technology growth be consistent with privacy and security. Also, the free expression of ideas must be protected, so technological development must proceed in a way that is consistent with an open Internet.

Challenges to Broad Adoption of Telecommunications Technology

Broad adoption of telecommunications technologies faces several challenges. For example, these technologies are unevenly adopted across different education and income levels. Home broadband adoption is more than twice as high for college graduates as for high school dropouts. Overall, 30 percent of Americans do not use broadband at home, and many of these non-users are in lower-income households. Rural areas also lag in adoption. As illustrated in Figure 5-9, nearly all urban residents have access to 6 megabits per second downloads, but only 82 percent of residents in rural communities can access those speeds, and the disparity becomes even larger at faster speeds.

One reason some households do not adopt broadband is cost: unlike the sharp price declines seen for technological hardware, such as computers, the prices consumers pay for Internet access have remained steady or risen. But while broadband prices have not fallen sharply, the speeds that are available at a given price today are often considerably faster than the available speeds at the same price several years ago, which means that value for money

Box 5-5: Electronic Health Records

Technological advances in Health Information Technology, especially Electronic Health Records (EHR), hold the promise of improving patient care and lowering health care costs. Patients are often treated by multiple providers for the same condition or for related conditions. Because the correct treatment by one provider often depends upon what other providers are doing, effective coordination of care between providers can improve health outcomes. Effective coordination also helps to control costs, as it avoids both the costs of treating follow-on problems resulting from uncoordinated care and the unnecessary duplication of tests and procedures.

Some ways of improving care coordination among health care providers involves changing the way they are paid for their services. The Affordable Care Act of 2010 (ACA) included a variety of such reforms that are currently in various stages of implementation, many of which are discussed in Chapter 4. But other ways involve the application of better technology, notably EHR systems. As the name would suggest, these systems enable the creation of a permanent, sharable record of every aspect of a patient's care, including test results, past treatments, and providers' notes. In a fully integrated EHR system, each provider has immediate and complete access to all relevant patient information, which has the potential to greatly improve coordination of care and also to reduce medical errors.

EHR systems have additional functionality as well, such as providing automatic alerts when treatments are inconsistent with each other or when a scheduled test has been missed. The systems can also be used to improve quality more broadly by allowing hospitals and other providers to keep better track of outcomes and to identify problem areas.

EHR adoption has been promoted by Administration policy. The Health Information Technology for Economic and Clinical Health (HITECH) Act, enacted as part of the American Recovery and Reinvestment Act of 2009, encouraged adoption and use of health information technology, including EHR systems.

Key programs established by the HITECH Act were the Medicare and Medicaid EHR programs. These programs provide financial incentives to hospitals and health care professionals to adopt EHR systems, and require that they demonstrate "meaningful use" of the systems. The meaningful use criteria, which become increasingly rigorous over time, require providers to demonstrate that they are using EHR systems to capture patient health information, assist in clinical decision making, track quality of care, and securely exchange patient information across health care settings to facilitate coordinated care.

Providers who adopt and demonstrate meaningful use of EHR systems by 2014 (for Medicare) and by 2016 (for Medicaid) are eligible for bonus payments from those programs. The Medicare program, but not the Medicaid program, also includes a payment reduction to providers that do not adopt and demonstrate meaningful use of EHR systems. Medicare providers who have not demonstrated meaningful use by 2015 are subject to penalties that grow over time; for example, for physicians, penalties start at 1 percent in 2015 and grow to 3 percent or more in subsequent years. The Congressional Budget Office estimated that the Medicaid EHR program would award bonuses of $12.7 billion through 2019, while the Medicare EHR program would make bonus payments net of penalties of $20 billion over that period (CBO 2009).

The HITECH Act also provided $2 billion to the Department of Health and Human Services to fund activities to encourage the diffusion of health information technology, such as investing in infrastructure and disseminating best practices. The Act also made a variety of other changes, including provisions to facilitate data sharing across health care providers to support coordinated care and protect patient privacy.

The share of medical providers using EHRs has risen dramatically in recent years. Data from the National Ambulatory Medical Care Survey show that the share of office-based physicians using an advanced EHR system (which are generally more sophisticated than those required meet the early-stage "meaningful use" criteria) rose from 17 percent in 2008 to 40 percent in 2012 (Hsiao and Hing 2014), and data from the American Hospital Association's annual survey of hospitals show that the share of hospitals that had adopted such a system rose from 9 percent to 44 percent over the same period (Charles et al., 2013). Consistent with this rapid progress for advanced systems, the Department of Health and Human Services has estimated that, as of the end of April 2013, over half of eligible physicians and more than 80 percent of eligible hospitals have adopted an EHR system and met the criteria for meaningful use (HHS 2013).

has improved. Further, while international comparisons are difficult (due to variations in factors like taxes, government subsidies, geography, population density, and product bundles), the United States compares favorably in a number of respects, including entry-level pricing for slower but still useful broadband speeds.

A surprisingly large number of households cite a different factor for their decision not to subscribe to home Internet service: a perceived lack of relevance to their day-to-day lives. Private- and public-sector broadband adoption programs address this by focusing on educating non-subscribers about the array of services and support mechanisms that are available online, like job listings and training, educational tools, health care services, and government resources.

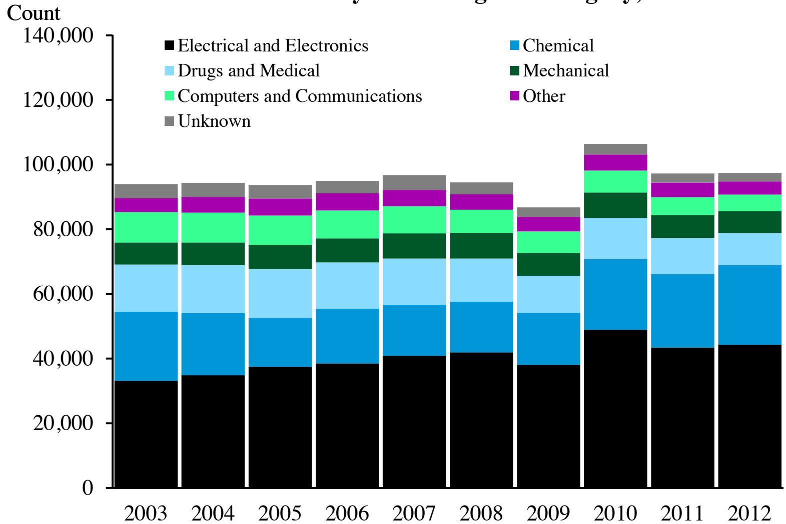
While this chapter has focused heavily on telecommunications technology, there are many other areas where technological advances promise large social and economic benefits, and where public policy can play an important role. One important example, discussed in Box 5-5, is Electronic Health Records and related technologies.

Patents

The rights that prospective innovators have to the economic returns on their innovations are known as intellectual property (IP) rights, of which one major category is patents, which apply to inventions. Patents are granted on inventions in many different areas of technology, as shown in Figure 5-10. The basic economic logic behind patent protection is simple: successful inventions are valuable to society, as they lead to new and better products. But attempting to invent is costly and risky. If successful inventions could be easily imitated by competitors, then prospective inventors may be in a position where they lose if their invention fails, but gain little or nothing if it succeeds. This diminishes the incentive to expend resources and effort on inventing, and the reduced rate of invention is harmful to society. To prevent this problem, patent protection allows inventors to enjoy a temporary exclusive right to their invention. The super-competitive pricing that results from this exclusivity provides an incentive to invest. Another benefit of patent protection is that patents are published, so they can be licensed and put to other socially valuable purposes other than those of the inventor. But patent protection can also harm consumers: for inventions that would have been created with weaker patent protection or even with no protection at all, patents simply lead to higher prices for the same inventions, not to additional inventions. The economically optimal strength of patent protection (for example, how many years a patent should run) is the one that best balances the benefit from accelerated invention with the harm from higher prices.

There are some additional effects of patent protection that also deserve mention. One effect is that some inventions are complementary to each other, meaning that the availability of one makes it easier to develop others. In those cases, the higher prices resulting from patent protection, as well as the related legal and administrative burdens (such as negotiating licenses), raise the cost of, and hence reduce the rate of, subsequent innovation. This effect is relevant for determining the economically optimal patent

Figure 5-10
Patents Issued in the U.S. by Technological Category, 2003–2012



Note: "Unknown" indicates that a patent could not be assigned to a single technological category with a high enough probability.
Source: United States Patent and Trademark Office.

strength. Another effect is that, in some cases, patent rights can be used to harm rival firms or to extract license fees that do not correspond to the value of the patent. As discussed below, it is important to curb such behaviors by developing sound policies related to patent examination and enforcement. It is also important to ensure that patents are not wrongly issued, but rather are only issued for inventions that are non-obvious, useful, and inventive.

The chapter now turns to two specific patent issues that have been the subject of recent policy scrutiny: how to support standard setting by appropriate use of standard-essential patents, and the activities of Patent Assertion Entities and the effects of those activities.

Standard-Essential Patents

We take for granted that we can drive our car up to a gas pump and have the hose fit the car's nozzle. Similarly, that smartphones created by different manufacturers will communicate with each other. These are examples of *interoperability* that result from the standardization of certain product features. An interesting problem arises when an industry seeks to adopt an interoperability standard and the available choices for the standard may include patented inventions.

The nature of the economic problem is to develop a mechanism that determines when standardization would make market participants better off and, in such cases, provides parties with incentives to invent and propose the invention as a standard, while ensuring that all parties will later find it in their interest to implement the standard. The central premise of the economic theory of patents is that granting limited exclusive rights through the issuing of patents provides an incentive for private investment in invention; absent such rights, the entity making the investment may not receive sufficient returns to make the investment worthwhile. These exclusive rights are not meant to preclude similar technologies from being developed and marketed. In principle, some degree of competition in consumer markets bounds the power conferred by these exclusive rights. But that bound is removed if a patented technology becomes the standard and is used in all products sold in the market. As a result, the patent holder may seek to charge higher prices than originally agreed on during the standard-setting process and to use the patent to inefficiently restrict access to the technology. Such behavior may delay implementation of the technology, as others who may adapt the technology exit the market or seek other ways around it. The sought-after standard then fails to become standard (for example, Gilbert 2011).

Because industry actors are most likely to understand the substantial complexities of new technologies and the potential products and markets for their dissemination, there is value in having the standards set voluntarily by industry-based standards-developing organizations (DOJ 2013). These organizations provide a place for industry actors to propose their patented technologies as part of a standard, and to reach consensus on the technologies incorporated into the standard (or to decide on no standard). After a decision is taken, a chosen patent becomes known as standard essential. Actual implementation of the agreed-on standard as an observed standard follows when all implementers and potential implementers pay an agreed justified price (reasonable, or both fair and reasonable) for the technology, and their access to the patented technology cannot be improperly restricted (there is not discrimination). By proposing a patent for use in the standard, the patent owner is giving up the power to charge higher per-unit prices for use of the technology, but enjoys returns from the diffusion of the technology more widely across more units.

Because the notion behind standard-developing organizations is voluntary collaboration, there is no guarantee that a standard will be produced. A standard-essential patent holder can refrain from committing to licensing on reasonable and non-discriminatory (RAND) terms.[14] In such cases, the declared standard is less likely to be the implemented standard and

[14] Sometimes the licensing commitment is to fair, reasonable and non-discriminatory terms.

market forces may be suggesting that a standard is not needed or may be best determined over time and in the marketplace directly (Farrell et al. 2007).

When voluntary agreement does produce a standard, there are instances when parties to the agreement do not feel that others are living up to the agreement. In such instances, when patent holders have committed to license on RAND terms, judicial and enforcement procedures should aim to reproduce the intent of the agreement; that is, to ensure that the patent holder receives a RAND royalty. Otherwise, judicial and enforcement procedures can tip the balance of power in favor of one party or the other, leading either to excessive market power in the hands of the patent holder or to non-payment of reasonable royalties by implementers, and to greater incentives against establishing a standard in the first place (Lemley and Shapiro 2005).

Patent Assertion Entities

In recent years, organizations known as Patent Assertion Entities (PAEs) have become common. PAEs brought 24 percent of all patent lawsuits in 2011, and over the 2007-11 period they brought approximately one-fifth of all patent lawsuits, covering about one-third of all defendants (GAO 2013). These PAEs purchase rights to patents belonging to other firms, and then assert them against firms or individuals who are using the patented technology. Some of this activity is valuable: incentives to invent are stronger if inventors know they can later sell their patent to, or merely engage the services of, a PAE that can assert it more effectively than they could do themselves. Also, in some cases, it may be efficient for PAEs to act as intermediaries by obtaining the rights to patents held by disparate inventors in order to decrease the transaction cost of negotiating licenses. However, many industry observers believe that PAEs often do not assert patents in good faith, but rather assert them simply in order to extract nuisance payments from firms looking to avoid costly and risky litigation. In some cases, these patents are valid but of low value, meaning that absent the high cost of litigation they would only command very low licensing fees. In other cases, the patents are invalid (or not infringed), and absent the high litigation costs they would not command any license fees at all (Scott Morton & Shapiro 2013).

This issue is particularly pronounced in smartphones and other consumer electronics devices (Chien 2012). Many of these products contain technology based on thousands of patents, and as shown in Figure 5-10 above, the number of patents issued in the "Electrical and Electronics" category has been increasing over the past decade. The large number of patents embodied in these products, and their complexity, often makes manufacturers subject to patent-infringement claims, with the associated threat of costly and risky litigation, from owners of low-value valid patents or even from

Box 5-6: The Leahy-Smith America Invents Act

The Leahy-Smith America Invents Act of 2011 took some important steps to update the U.S. patent system. The Act changed the system to give priority to the first inventor to file for a patent on a given invention, moving away from the "first-to-invent" priority system and bringing the United States in line with every other industrialized nation. This change eliminated the need for long, expensive administrative proceedings to resolve disputes among inventors who filed for patents on the same invention over who invented it first.

The Act also helps ensure that inventors have the opportunity to share their work early on by maintaining a form of the one-year "prior art" grace period that had been a feature of the previous system. The grace period excludes from the previous state of knowledge, against which the originality of a patent application is judged, any disclosures of details of the invention made within the year preceding the application date by the inventor, or by third parties who learned them from the inventor. The grace period allows inventors to publish their work, prepare application materials, or seek to raise funds to support their application without fear that those activities will later be a detriment to that application.

The Act also increases protections from patent infringement lawsuits for innovators who develop and deploy new products or methods but choose not to patent them, a common practice in the high-tech industry, by expanding the "prior user rights" infringement defense. Formerly applicable only to business practices patents, this defense—which exempts from liability users who can demonstrate that they independently developed and used the patented product or method that they are accused of infringing upon, and did so more than a year prior to the date the patent was filed—is now applicable to all types of patents.

owners of invalid patents. It is therefore an important public policy goal to find ways to reduce the cost of defending patent lawsuits, and also to reduce the number of invalid patents, either by reducing the number of invalid patents that are granted, or by making it easier for them to be challenged.

One important step toward resolving these patent-related problems, which disrupt the appropriate economic incentives to invent, has been taken in the form of the Leahy-Smith America Invents Act of 2011, discussed further in Box 5-6. The key provisions of the AIA, which went into full effect in 2012, are helping to improve the patent system for innovators by offering a fast-track option for patent processing, taking important steps to reduce the current patent backlog, and increasing the ability of Americans to protect their intellectual property abroad.

Several provisions of the America Invents Act may help address some of the problematic behavior of PAEs by developing at the Patent and Trademark Office new programs to create alternatives to litigation over patent validity, new methods for post-grant review of issued patents, and major steps to increase patent quality through clarifying and tightening standards. Yet challenges remain, notably the asymmetry between the cost to a PAE of bringing a patent lawsuit and the cost to a target firm of defending one, which enables PAEs to bring weak cases in the hopes of extracting a settlement.

In June 2013, the President issued a set of five executive actions and seven legislative recommendations to address these challenges. These included measures to make it more difficult for overly broad claims to receive patents in the first place, as well as to make it easier to challenge weak patents once they have been granted. The President's priorities also include measures to require greater clarity in patent applications regarding the precise nature of the claimed invention, as well as the identity of the patent holder. Other measures include ways to make it more difficult for patent holders to sue end-users (as opposed to manufacturers) of products that contain patented technology, and to provide judges with more discretion to award attorney fees and other costs to the prevailing party in patent lawsuits.

Congress has taken up these issues. In December 2013, the House of Representatives passed a bipartisan bill containing many of the Administration's priority items. A related bill is currently under consideration in the Senate Judiciary Committee.

Another important policy issue related to patents is the phenomenon of "pay-for-delay" settlements of patent lawsuits in the pharmaceutical industry. This is discussed in Box 5-7 below.

Conclusion

Productivity growth allows a given set of scarce resources to yield more output and a higher aggregate standard of living. When private actors face incentives that lead them to optimal investments in growth-enhancing technologies, government policy should be to not interfere. But at other times, a light touch from government is needed to align incentives or to act in place of incentives that are missing: in the form of conducting of its own research; or of subsidization of private research; or through appropriate intellectual property rights laws, regulation, and enforcement. Government also has a role to play in ensuring that all citizens benefit from productivity advances that can increase living standards—a step that can form a virtuous cycle that also increases productivity growth itself by tapping more of the potential of our citizens.

Box 5-7: Pay-For-Delay Settlements in Pharmaceutical Patent Cases

Out-of-court settlements of lawsuits are usually socially beneficial, as they allow disputes to be resolved without a costly trial. There are circumstances, however, where settlement of a patent lawsuit can be used as a means of extending market power, rather than as a means of efficiently resolving the dispute. In recent years, this has been a significant issue for certain cases involving pharmaceutical patents. In these cases, an incumbent seller of a branded drug files a patent infringement suit against one or more companies seeking Food and Drug Administration approval to sell a generic version of that drug. The patent at issue is often not the one covering the drug's active ingredient (for which assessing infringement is usually less complicated), but is instead a secondary patent, such as one covering a particular formulation of the drug (Hemphill and Sampat 2011). The generic entrant will deny the infringement, claiming that the patent is invalid, that its product does not infringe, or both. A patent lawsuit results, and is settled through an agreement specifying a date on which the generic entrant may begin selling its product, which is some time after the date of the settlement and before the patent expires. The settlement will also specify a payment from the branded incumbent to the generic entrant. Absent the settlement, the case would have gone to trial; had the incumbent won, the generic entrant would have been barred from entry until the end of the incumbent's patent term, and had the entrant won, it could have entered immediately and sold its product, assuming it had received FDA approval.

The willingness of the incumbent to agree to such a settlement may seem puzzling, as the payment appears to go the "wrong" way, from the alleged infringer to the infringed. But the ability to enter into such settlements can benefit the incumbent by enabling it to "purchase" later generic entry than would otherwise occur. In other words, settlements of patent disputes can be used as a vehicle for extending market power.

What drives these settlements is the fundamental economic principle that the profits of a single seller of a product are greater than the combined profits of two or more sellers, because a single seller has greater market power and so can extract a higher price from consumers. A settlement that delays generic entry of a drug therefore increases the aggregate profits on that drug. These extra profits create an incentive for a deal in which entry is delayed; both parties will accept such a deal as long as the extra profits are divided in such a way that each party is better off than it would be absent the deal (i.e., better off than by letting the patent lawsuit proceed to trial). For this reason, these settlements are often called "pay-for-delay" settlements.

Pay-for-delay settlements undermine existing laws (most notably the Hatch-Waxman Act) that encourage the development of generic drugs. When generic drugs enter a market, they are offered at a much lower price than is the branded drug, and they typically capture a large market share. For these reasons, generic entry results in considerable savings to consumers and to the health care system. The delay of generic entry due to pay-for-delay settlements greatly reduces those savings.

The ability of incumbent patent holders to enter into pay-for-delay settlements, and to thereby maintain their patent protection for a longer period, might be viewed as increasing the value of pharmaceutical patents, and hence increasing the incentive to invest in discovering new drugs. However, the value of any increased innovation arising from these settlements may be relatively small. The most socially valuable drug patents are often those covering new molecular entities. These patents are relatively unlikely to be successfully challenged, which means the generic entrant has little prospect of victory at trial or of a lucrative pay-for-delay settlement. As a result, banning such settlements may not significantly affect the incentive to invest in inventing new molecular entities. Instead, pay-for-delay settlements often involve patents on incremental improvements to existing drugs, often ones that make the drug just different enough that a prescription for the new version cannot be filled with an existing generic equivalent of the old version. The ability to enter into pay-for-delay settlements does encourage this type of innovation, but the social benefits are likely to be comparatively small in many cases.

Pay-for-delay settlements have been the subject of a considerable amount of litigation, culminating in a 2013 Supreme Court decision in FTC v. Actavis, involving a drug called AndroGel. The Court ruled that "pay for delay" settlements are not presumptively unlawful, but are also not immune from antitrust scrutiny, partially resolving earlier conflicting rulings by lower courts (see FTC v. Actavis 2013). The Court did not establish a concrete rule regarding how such settlements should be treated, however, so substantial uncertainty remains about how these lawsuits will be adjudicated in practice.

The Administration has proposed legislation that gives the Federal Trade Commission explicit authority to stop companies from entering into pay-for-delay agreements. For the reasons described above, such authority would likely generate billions of dollars in savings for consumers, and also for the Federal Government through lower pharmaceutical prices paid by Medicare, Medicaid, the Department of Defense, and the Veterans Administration (see CBO 2011 and FTC 2010).