Introduction

The effect of technical progress on the welfare of workers has long been a matter of controversy. Historically, one can document famous episodes of violent protests against productivity improvements that workers felt threatened their jobs. In *Das Kapital*, Marx (1867) documents an episode of revolt against the introduction of machinery, which actually led to innovation being stalled:[1]

> In the 17th century nearly all Europe experienced revolts of the workpeople against the ribbon-loom, a machine for weaving ribbons and trimmings, called in Germany Bandmühle, Schnurmühle, and Mühlenstuhl. These machines were invented in Germany. Abbé Lancellotti, in a work that appeared in Venice in 1636, but which was written in 1579, says as follows: "Anthony Müller of Danzig saw about 50 years ago in that town, a very ingenious machine, which weaves 4 to 6 pieces at once. But the Mayor being apprehensive that this invention might throw a large number of workmen on the streets, caused the inventor to be secretly strangled or drowned."
>
> In Leyden, this machine was not used till 1629; there the riots of the ribbon-weavers at length compelled the Town Council to prohibit it.

In 1768, a group of spinners broke into the home of James Hargreaves, the inventor of the "Spinning Jenny," a machine which was capable of doing the work of eight workers, and destroyed his machines.

In the early nineteenth century, textile workers—the Luddites—organized against the introduction of advanced machinery that made their skills redundant. This movement is described as follows on the Web page[2] of Dr. Steve Anderson from Utah University:

> For at least three hundred years the weavers from in and around the central English town of Nottingham, though commoners, enjoyed the status and rewards accorded to fine craftsmen. The weavers of Nottinghamshire produced lace and stockings that dominated the English markets and were prominent items in export trade. These products were hand made, often in the weaver's home.... In the first years of the 19th century stocking frames and the early automation of the power loom threatened this long-standing way of life.... The weavers complained bitterly that the machines made mass produced products of shamefully inferior quality. Naturally, the weavers saw the new technology

[1] See www.marxists.org/archive/marx/works/1867-c1/ch15.htm#S5.
[2] See www.usu.edu/sanderso/multinet/lud1.html.

> as the most powerful tool of their new oppressor, the factory owner.... During a short period climaxing in the spring of 1812, inspired perhaps by the French Revolution and the writings of Thomas Paine, the weavers formed into something akin to a guerrilla army and took substantial control over the territory near Nottingham and several neighboring districts.... The Luddites often appeared at a factory in disguise and stated that they had come upon the orders of General Ned Ludd. These demands included restoration of reasonable rates of compensation, acceptable work conditions, and probably quality control. Faced by the intimidating numbers and the surprisingly disciplined actions of the Luddites, most factory owners complied, at least temporarily. Those that refused found their expensive machines wrecked.... The nonviolent period of Luddism ended at Burton's power loom mill in Lancashire on April 20, 1812. A large body of Luddites, perhaps numbering over a thousand attacked the mill, mostly with stick and rocks.... A government crackdown ensued, and many suspected Luddites were convicted, imprisoned, or hanged.

Such incidents have led to the famous controversy between Marx and Ricardo over the role of machinery. Marx forecast a world where innovation made workers ever more useless, leading to their impoverishment, along with a secular increase in the share of capital in national income. On the other hand, Ricardo and the neoclassical economists who followed him thought that innovation allowed a single worker to produce more output per unit of time, which led to an increase in wages and living standards.[3]

The explosion in living standards over the last two centuries has proved that the neoclassicists were right, while Marx was wrong. This is why the most influential growth model used by economists is the Solow (1956) one, where the economy converges to a balanced growth path in which wages grow in line with productivity. In recent years, however, economists have documented a worrying trend toward greater wage inequality in the United States and other countries. Not only has the distribution of wages widened, but real wages have fallen for the lowest paid workers (the bottom 20%, say), despite continuing growth in GDP per capita.

A large empirical literature has studied this phenomenon[4] and has found that the returns have increased for all the dimensions of skill:

[3] The neoclassicists in fact admitted a negative impact of technical progress on wages in the short run, but not in the long run. They had in mind the short-run complementarities between labor and capital that are analyzed in chapter 2 of this book (see Beach 1971).

[4] See in particular Juhn et al. (1993), Levy and Murnane (1992), and Katz and Murphy (1992). An important survey, which in many ways complements this book, can be found in Acemoglu (2002b).

education, experience, and unobserved ability. A number of explanations have been proposed and they are detailed in the following paragraphs.

A first explanation, put forward by, for example, DiNardo et al. (1996) and Blau and Kahn (1996), ascribes the rise in inequality to an increased role for market forces, relative to institutional forces, in the determination of wages. As conservative governments came into power in the United Kingdom and the United States in 1979 and 1981, respectively, wages became more closely aligned with individual productivity, and were less determined by union contracts. This situation creates a move toward greater wage inequality because unions tend to compress wages between skill levels. On the other hand, unions tend to increase income inequality by putting some workers out of jobs, but that is not reflected in measures of wage inequality since these measures take into account only the employed. The merit of this explanation is that it accounts for the fact that inequality has not increased in countries such as France and Germany where labor-market institutions have not evolved. On the other hand, it fails to explain the fact that inequality started increasing around 1975, long before the reforms of Margaret Thatcher and Ronald Reagan were implemented.

A second explanation is that because of immigration by unskilled workers, skilled workers have become relatively scarcer in the labor market. This explanation has been discarded on the grounds that whereas immigrants tend to be less skilled than natives in the destination country, this trend is more than offset by the increase in secondary and tertiary education enrollment rates, which tends to raise the relative supply of skilled workers.[5] Thus, it appears that the rise in wage inequality must be explained by shifts in relative demand, rather than shifts in relative supply.

A third explanation is international trade. This says that developed countries are now immersed in a world economy in which factor-price equalization prevails. The relative wages of unskilled workers are now determined by their relative scarcity worldwide, rather than within a given country. As a result, integration in the world economy should be associated with widening wage inequality in developed countries, and shrinking inequalities in developing countries. This is essentially the famous Stolper–Samuelson theorem (see Stolper and Samuelson 1941): any factor that is scarce in a given country, relative to the rest of the world, sees its return fall when the country opens up to trade. While the debate on this hypothesis is not totally settled, Lawrence and Slaughter

[5] Note that a similar point can be made with respect to the increase in female labor-market participation. On the effect on labor-market outcomes, see Borjas et al. (1997).

(1993) have shown that it is inconsistent with a number of basic facts. Most notably, if international trade were the driving force for the rise in inequality, we should observe that within each firm in the West the ratio of unskilled workers to skilled workers increases. The argument is that if technology has not changed, firms are faced with a lower relative wage of the unskilled, which induces them to increase their relative input of that factor. Of course, at the aggregate level, activity is relocated away from unskilled-intensive industries toward the skilled intensive ones; the unskilled-intensive industries are faced with a fall in their prices, because of competition from less developed countries, and therefore reduce their activity. However, it remains true that within any *existing* firm, one should observe a greater use of unskilled workers. Indeed, the reallocation of activity toward skilled-intensive industries is the countervailing force necessary to keep the skill ratio unchanged at the aggregate level. Lawrence and Slaughter have shown that the skill ratio did not evolve in favor of unskilled workers at the firm level; on the contrary, firms increased their demand for skilled workers.

This suggests that there has been technological change that has altered the relative factor demand in favor of skilled workers—the fourth explanation.[6] A key potential driving force has been the sharp progress in information technology, which picked up in the mid 1970s. Research has documented how computers and a number of other information and communication technologies (ICTs) are complementary with skilled labor and substitute for unskilled labor.[7]

While the empirical research has been quite extensive, this book's ambition is to provide an analytical perspective on the effects of technical progress on the distribution of income. I hope to show that these effects are pervasive, and that one cannot simply think of "information technology" as a black box that just shifts the relative demand of one category of workers.

---

[6] The theoretical literature has also investigated the interactions between various explanations. For example, Acemoglu et al. (2001) study a model in which skill-biased technical progress induces a fall in the power of unions, because skilled workers opt out from them as their outside option in bargaining has gone up. There also exists a burgeoning literature on the interactions between trade and technology. For example, Grossman and Maggi (2000) use the class of models described in chapter 8 of this book to derive novel predictions about how differences in the *distribution* of skills between countries with identical aggregate factor endowments and identical technologies lead to gains from trade. Manasse and Turrini (2001) use the superstars models analyzed in our chapter 7 to argue that globalization leads to an increase in market size for the most talented workers, which tends to widen the distribution of income. Thoenig and Verdier (2003) argue that skill-biased technical change may arise as protection against imperfect enforcement of intellectual property abroad.

[7] See in particular Autor et al. (1998, 2003).

This book studies how different categories of workers gain or lose from innovation, and it shows that this crucially depends on the nature of innovation. In particular, one must consider the following questions:

- Does innovation increase the efficiency of physical capital, human capital, or labor?
- Does innovation affect the way workers interact with each other, for example by making it easier to communicate and to access knowledge, or by changing the number of people affected by a given worker's activity?
- Does innovation affect the way the economy is organized? For example, how is the hierarchical structure of the workplace changed? Is the allocation of resources to different stages of production affected?
- Is innovation "vertical" (increasing productivity in existing goods) or "horizontal" (increasing the range of available goods)?

Another important methodological aspect is how we think of the labor market. In macroeconomics, one typically assumes that labor can be reduced to a vector of homogeneous inputs in the production function. Thus, production depends on the aggregate input of raw labor, human capital, and perhaps more detailed things such as beauty, strength, health, etc. This means that an individual endowed with, say, 0.6 units of strength and 0.4 units of intelligence would earn $0.6w_S + 0.4w_I$, where w_S is the market price of strength and w_I is the market price of intelligence. Another individual endowed with 1 unit of strength and 0.7 units of intelligence would earn $w_S + 0.7w_I$. Furthermore, the first individual might just decide to save on his or her strength and intelligence and just supply half of them, then earning $0.3w_S + 0.2w_I$. This is what we can call the "homogeneous-input" view of the labor market. Many labor markets do not work in this way, in that individuals are actually unique and not reducible to a fungible vector of characteristics. One cannot be a part-time professional footballer, for example. Nor can one replace a good professional footballer with two bad ones. And an intelligent footballer cannot supply his football skills to Manchester United and his intelligence to MacKinsey at the same time. This can be referred to as the "quality-input" view of the labor market.

This book shows that the effect of technology on the distribution of income differs greatly according to which of these two views of the labor market one takes.

Under the homogeneous-input view, what is important is the degree of substitutability and complementarity between technology and the different inputs in the production function. Technology harms substitute factors and benefits complementary factors. In the long run, additional effects arise from the fact that some factors (such as physical and human capital) can be accumulated. For example, technology may be detrimental to labor and beneficial to capital if it is substitutable with labor, i.e., if it is equivalent to an increase in the labor input. But this increases the return to capital, which induces people to accumulate more capital, which in turn boosts wages. One can show that under standard (but not necessarily correct) assumptions, wages cannot then fall in the long run.

Under the quality-input view, a key aspect is the extent to which high-quality individuals can spread their talent over a larger market. For a footballer, this might mean having more viewers thanks to satellite television. For a manager, it might mean supervising more people by means of e-mail and video conferencing. For an accounting consultant, it might mean embodying his knowledge in a piece of software that can be distributed to millions of clients. A recurring theme is that technical progress does not then harm the poorest, because they are not substitutes for quality workers. Rather, they can watch better football, are supervised by better managers, and are more productive because their firm, thanks to the new software, accesses better accounting services at a lower cost. Those who suffer are those who are at the bottom of the ability distribution of quality workers, say the average accountant who cannot compete with the new software and is displaced to another occupation—presumably one where individual quality is less important and where the labor input is more homogeneous.

Chapters 2–5 discuss the impact of technology on wages if labor inputs are homogeneous. I start with simple neoclassical growth models, comparing the short-run effects with the long-run effects, and then move to more complex models with three factors of production: labor, human capital, and physical capital. I also distinguish between technical change as a shift to a single production function for the economy and technical change as the introduction of a new technology, summarized by an alternative production function, which may be used along with the old technology. I discuss the possibility of counterintuitive *supply effects*: for instance, the fact that an increase in the supply of human capital may trigger a rise, rather than a fall, in its return, because more use is made of technologies intensive in that factor.

Chapters 6–8 analyze the impact of technical change when labor is a quality input. I focus on two key phenomena. First, more talented individuals can spread their talent over a larger market—"The Economics of

Superstars." Technical change then affects the relative "market size" for two individuals with different skills, which in turn has an effect on the distribution of income. Second, a worker's quality affects the productivity of his or her coworkers. These interactions within the firm determine its willingness to pay for workers of different quality. It also affects the way workers of different quality are matched together. Depending on the technology, workers of similar qualities are matched in the same firms—a segregated outcome—or firms hire workers with heterogeneous skills—a "unitary" outcome. Technical change has an effect on both the nature and the size of the interactions between the workers, and then again on both the pattern of segregation and the distribution of income.

In all of these examples, technical change has a nonproportional impact on the marginal product of different factors of production. It may come as a surprise that the distribution of income may also be changed when technical progress has a proportional effect on marginal products. Yet this is possible, through general equilibrium demand linkages, if there are nonproportionalities in demand. Chapters 9 and 10 study these nonproportionalities. One can show that growth can affect the distribution of income through its effect on monopoly markups if the utility of consumers is nonhomothetic. This evolution of markups may be immiserizing for workers if utility is such that the price elasticity of demand for a good falls when consumers become richer. This will be the case, in particular, if the level of utility that can be obtained from consuming a good is bounded (an assumption weaker than, but similar to, that of satiation). Roughly speaking, the rich become "careless" consumers as they near satiation, which triggers a rise in monopoly markups and harms workers.

Another interesting consequence of nonhomothetic utilities is that conflicts of interest between the rich and the poor arise with respect to the level of innovation: under plausible conditions, the rich value the introduction of new goods more than the poor do.

We also study, building on work by Matsuyama (2002), how technical change affects the distribution of welfare through its effect on the range of goods consumed by different people. It is shown that, under a utility function with a hierarchy of needs, technical change may trigger a virtuous circle: the economy converges to a "mass consumption society." As new consumers of industrial goods start buying them, those goods can be produced more efficiently, because of learning by doing. This reduces the cost of these goods, inducing more people to buy them. These new consumers in turn lead to further advances in productivity, and so on.

This book is based on lectures that I have given to Ph.D. students in Toulouse, at MIT, and in Paris, as well as to students in summer schools in

Barcelona, Kiel, Rotterdam, Steyr, Ammersee, and Gerzensee.[8] It requires technical knowledge of economics at M.Sc. level, and the reader with such knowledge can start the book from chapter 2. On the other hand, the book should be accessible to readers with less knowledge of economics but basic proficiency in calculus and analysis, provided they read the introductory material presented in chapter 1.

[8] It has benefited from their comments, as well as comments by David Autor, Richard Baggaley, and two anonymous reviewers.