NBER WORKING PAPER SERIES

TECHNICAL CHANGE, INEQUALITY, AND THE LABOR MARKET

Daron Acemoglu

Working Paper 7800
http://www.nber.org/papers/w7800

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2000

I thank Alexis Leon for excellent research assistance and Joshua Angrist, David Autor, Olivier Blanchard, Lawrence Katz, Omer Moav, and Steve Pischke for comments. The views expressed here are the author and do not necessarily reflect those of the National Bureau of Economic Research.

Technical Change, Inequality, and The Labor Market
Daron Acemoglu
NBER Working Paper No. 7800
July 2000
JEL No. J30, J31, O14, O31, O33

ABSTRACT

This essay discusses the effect of technical change on wage inequality. I argue that the behavior of wages and returns to schooling indicates that technical change has been skill-biased during the past sixty years. Furthermore, the recent increase in inequality is most likely due to an acceleration in skill bias. In contrast to twentieth century developments, most technical change during the nineteenth century appears to be skill-replacing. I suggest that this is because the increased supply of unskilled workers in the English cities made the introduction of these technologies profitable. On the other hand, the twentieth-century has been characterized by skill-biased technical change because the rapid increase in the supply of skilled workers has induced the development of skill-complementary technologies. The recent acceleration in skill bias is in turn likely to have been a response to the acceleration in the supply of skills during the past several decades.

Daron Acemoglu
Department of Economics
MIT, E52-371
50 Memorial Drive
Cambridge, MA 02142-1347
and NBER
daron@mit.edu

1 Introduction

1.1 Motivation

What are the implications of technical change for the labor market? How does new technology affect the distribution of wages and income? Is technology responsible for the changes in the wage structure observed in many advanced economies since the 1970s?

The recent consensus is that technical change favors more skilled workers, replaces tasks previously performed by the unskilled, and exacerbates inequality. This view is shaped largely by the experience of the past several decades, which witnessed both major changes in technology, including the rapid spread of computers in workplaces and in our lives, and a sharp increase in wage inequality. In the U.S., for example, the college premium—the wages of college graduates relative to the wages of high school graduates— increased by over 25 percent between 1979 and 1995. Overall earnings inequality also increased sharply. In 1971, a worker at the 90th percentile of the wage distribution earned 266 percent more than a worker at the 10th percentile. By 1995 this number had risen to 366 percent (author's calculations from March CPS data). Many commentators see a direct causal relationship between technological changes and these radical shifts in the distribution of wages taking place in the U.S. economy. The title of Krueger's (1993) influential paper on computers and inequality summarizes this view: *"How Computers Have Changed the Wage Structure."* Greenwood and Yorukoglu (1997, p. 87) similarly give a succinct statement:

> "Setting up, and operating, new technologies often involves acquiring and processing information. Skill facilitates this adoption process. Therefore, times of rapid technological advancement should be associated with a rise in the return to skill."

They further argue that we are now in the midst a "Third Industrial Revolution", fueled by advances in information technology, and that this revolution is responsible for the increase in inequality (as does Caselli, 1999, in a paper entitled "*Technological Revolutions*").

The view that technological developments favor skilled workers receives support from accounts of earlier episodes. For example, there were already signs of significant *technology-skill complementarity* in the 1910s. Goldin and Katz (1998) argue that the spread of batch and continuous-process methods of production increased the demand for skills. They add "...the switch to electricity from steam and water-power energy sources was reinforcing because it reduced the demand for unskilled manual workers in many hauling, conveying, and assembly tasks." (p. 695). Over this period, capital-intensive industries increased the demand for skills considerably (see Goldin and Katz, 1998, Table 3), and the scope of these industries expanded with the sharp fall in the price of electricity (see, for example, Woolf, 1984, p. 178). The rapid increase in the importance of white collar and clerical occupations gave another boost to the demand for skills. Generalizing from the experience of the 1920s, Harry Jerome (1934, p. 402) argued that "...in the future...there is considerable reason to believe that the effect of further [mechanization] will be to raise the average skill required."

The early twentieth century evidence was so powerful that Griliches (1969) suggested capital and skills are intrinsically complementary. Nelson and Phelps (1967), Welch (1970), Schultz (1975) and Tinbergen (1975) also argued that technological developments increase the demand for skills. Events since then support this notion. Personal computers, computer-assisted production techniques and robotics appear to complement skilled workers, replacing many labor intensive tasks. In this light, it is perhaps natural to view the increase in inequality over the past several decades as a direct consequence of technical change.

Although the consensus is now broad, the idea that technological advances favor more skilled workers is a twentieth century phenomenon. In nineteenth century Britain, skilled artisans destroyed weaving, spinning and threshing machines during the Luddite and Captain Swing riots, in the belief that the new machines would make their skills redundant. They were right: the artisan shop was replaced by the factory and later by interchangeable parts and the assembly line (e.g., James and Skinner, 1985, Goldin and Katz, 1998). Products previously manufactured by skilled artisans started to be produced in factories by workers with relatively few skills, and many previously complex tasks were simplified, reducing the demand for skilled workers.[1] Mokyr (1990, p. 137) describes this process vividly:

> "First in firearms, then in clocks, pumps, locks, mechanical reapers, typewriters, sewing machines, and eventually in engines and bicycles, interchangeable parts technology proved superior and replaced the skilled artisans working with chisel and file."

Interchangeable parts were in fact very much designed to be skill-replacing. Eli Whitney, a pioneer of interchangeable parts, described the objective of this technology as

> "to substitute correct and effective operations of machinery for the skill of the artist which is acquired only by long practice and experience; a species of skill which is not possessed in this country to any considerable extent." (quoted in Habakkuk, p. 22)

[1] It can be argued that technical change always increases the demand for "skills", and the artisans who were hurt as a result of new technology were not "skilled" since they lacked the flexibility to adapt to the required changes. This argument is not totally convincing, since the artisans earned considerably more than other laborers (for example, James and Skinner, 1985, report over 60 percent wage differentials for building and printing workers relative to laborers in the 1850s). So the artisans possessed skills that were being rewarded by the market, and the standardization of the production process destroyed these rewards. On the other hand, it has to be noted that many of the skill-replacing technologies of the nineteenth century may have also increased the demand for engineers and managers (see, e.g., Goldin and Katz, 1998).

The experience of the nineteenth and early twentieth centuries led Braverman (1974) and Marglin (1974) to argue that technical change was "deskilling"—a major purpose of technical change was to expand the division of labor and simplify tasks previously performed by artisans by breaking them into smaller, less skill-requiring pieces. Braverman (1974, p. 113), for example, suggested that the first principle of management and production techniques of the period was "dissociation of the labor process from skills of the workers. The labor process is to be rendered independent of craft, tradition, and the workers' knowledge."

A longer view therefore suggests that technological advances do not always increase the demand for skills. In fact, most nineteenth century innovations appear to have replaced skilled workers and expanded tasks performed by the unskilled. But then, why have technological advances been skill-biased in the twentieth century? And, are technological changes *the major cause* of the recent increase in inequality?

This essay attempts to answer these questions. It has two main theses:

1. The behavior of wages and returns to schooling indicates that technical change has been skill-biased during the past sixty years, and probably for most of the twentieth century. Furthermore, an acceleration in skill bias during the past few decades is the main cause of the increase in inequality.

2. We can understand the behavior of technical change by recognizing that the development and use of technology is, at least in part, a response to profit incentives.[2] When developing skill-biased techniques is more profitable, new technology will tend to be skill-biased. I suggest that the nineteenth century was characterized by skill-replacing developments because the increased supply of unskilled workers in the English cities (resulting from migration from rural areas and from Ireland) made the introduction of these technologies profitable. In contrast, the twentieth century has been characterized by skill-biased technical change because the rapid increase in the supply of skilled workers has *induced* the development of skill-complementary technologies. The recent more rapid skill-biased technical change is in turn likely to have been a response to the acceleration in the supply of skills during the past several decades. However, I also argue that despite the acceleration in skill bias, we are most likely not in the midst of a "Technological Revolution"; what has changed is not necessarily the overall rate of progress, but the types of technologies that are being developed.

[2]Precedents of this approach include Schmookler (1966), who emphasized *demand pull* and the extent of the market as key determinants of innovations; the endogenous growth theory, e.g., Romer (1990), Grossman and Helpman (1991), and Aghion and Howitt (1992); the induced innovation theory, including Ahmad (1965), Kennedy (1964), Samuelson (1970), Hayami and Ruttan (1970), and David (1975); and recent work including my own, Acemoglu (1998, 1999b, 2000), Acemoglu and Zilibotti (1999), and Kiley (1999).

Finally, I conjecture that recent technological developments are likely to have affected the organization of the labor market—including the way firms are organized and the form of labor market institutions— and may have had a large effect on the structure of wages through this channel.

In the process of developing this argument, this essay sets out a simple theoretical framework, in which inequality and returns to skills are determined by supply and demand forces (technology).[3] Using this framework as a unifying device, I critically survey many of the theories that explain the recent increase in inequality by technological factors, and discuss how various pieces of evidence can be interpreted within this framework.

1.2 Summary of the Argument

I begin with a roadmap of the argument. Since the effect of new technology on the distribution of wages in the recent past is central to the focus here, I organize this essay around a number of salient facts from the post-war U.S. economy.[4] Briefly, these facts are:

1. The past sixty years have seen a large increase in the supply of more educated workers, while returns to education have risen.

2. Returns to education fell during the 1970s, when there was a very sharp increase in the supply of educated workers. Returns to education then began a steep rise during the 1980s.

3. Overall wage inequality rose sharply beginning in the 1970s. Increases in within group (residual) inequality—i.e., increases in inequality among observationally equivalent workers— account for much of this rise.

4. Average wages have stagnated and wages of low-skill workers have fallen in real terms since 1970.

I argue that technical change over the past sixty years, or even over the past century, has been skill-biased. This conclusion follows from fact 1 above: in the absence of substantial skill bias in technology, the large increase in the supply of skilled workers would have depressed the skill premium. In 1970, Welch (1970, p. 36) reached the same conclusion, and argued:

[3] Precedents of the supply and demand approach include, among others, Becker (1964), Welch (1970) and Tinbergen (1975).

[4] I limit the discussion of the major trends to the U.S. economy because of space constraints, and also because there is notably more research to build upon.

> "With the phenomenal rise in average education, why have rates of return failed to decline?...
>
> It is obvious that changes have occurred to prevent the decline in returns to acquiring education that would normally accompany a rise in average educational level. Presumably, these changes have resulted in growth in demand for ... education... sufficient to absorb the increased supply with constant or rising returns."

The 30 years after Welch wrote these words witnessed a much more rapid increase in the supply of education, and a sharp increase in the returns to more skilled workers, suggesting that skill-biased changes in technology continued throughout the postwar period.

And yet, if technical change has been skill-biased throughout the recent past, why did inequality increase during the past 30 years, but not before? There are at least two possible answers to this question. The first, which I call the *steady-demand hypothesis*, maintains that demand for skills increases at a constant pace, so changes in inequality must be explained by the pace of the increase in the supply of skills. According to this hypothesis, inequality was relatively stable before the 1970s, because the rate of skill accumulation in the U.S. economy was more rapid than the constant pace of skill-biased technical change (e.g., Katz and Murphy, 1992). The recent increase inequality is then explained not by a major technological change, but by a relative slowdown in skill accumulation. The second possible answer comes from the *acceleration hypothesis*, which maintains that there has been an acceleration in skill bias beginning in the 1970s or 1980s. According to this hypothesis, there has been a notable acceleration in the demand for skills, driven in large part by advances in information technology, and perhaps even approaching the scale of a "Third Industrial Revolution".

So was there an acceleration in skill bias? This question is difficult to answer as we lack direct measures of the degree of skill bias of technologies. To tackle this question, one therefore needs to look at a variety of evidence often pointing in different directions. I conclude below that skill-biased technical change is likely to have accelerated over the past several decades. This conclusion is based on the sharp increase in overall inequality starting in the 1970s and on the fact that returns to schooling rose over the past thirty years despite the unusually rapid increase in the supply of educated workers.

Why did the demand for skills accelerate over this period? And why has new technology favored more skilled workers during the twentieth century, but not during the nineteenth century? One approach would view technology as exogenous, stemming from advances in science or from the behavior of entrepreneurs driven by a variety of nonprofit motives. Demand for skills increased faster during the past thirty years, this approach would maintain, because of a *technological revolution* led by the microchip, personal computers and the Internet.[5] New technologies of the nineteenth century were not skill-biased because the technological frontier then only enabled the invention of skill-replacing techniques.

Yet, there are a number of problems with this approach. First, although a number of papers, including Greenwood and Yorukoglu (1997), Hornstein and Krusell (1997), and Galor and Moav (2000), show that rapid technical change may lead to slower total factor productivity (TFP) growth, the slow rates of TFP and output growth of the past several decades are difficult to reconcile with a technological revolution during this time period. Second, demand for skills appears to have accelerated starting in the late 1970s, precisely when the supply of skills increased very rapidly. Exogenous technology theories do not explain the timing of this acceleration.[6]

An alternative theory maintains instead that new technologies are *endogenous* and respond to incentives. It was the large increase in the supply of skilled workers, this approach claims, that induced the acceleration in the demand for skills. The reasoning is as follows. When skill-biased techniques are more profitable, firms will have greater incentives to develop and adopt such techniques. A key determinant of the profitability of new technologies is their market size; machines that can be sold in greater numbers will be more profitable. Schmookler (1966), in his pioneering study, *Invention and Economic Growth,* placed great emphasis on market size. He argued (p. 206) "invention is largely an economic activity which, like other economic activities, is pursued for gain;... expected gain varies with expected sales of goods embodying the invention." This reasoning implies that machines complementary to skilled workers will be more profitable to develop when there are more skilled workers to use them. New technologies have become more skill-biased throughout most of the twentieth century because the supply of skilled workers has grown steadily. This perspective also suggests that a faster increase in the supply of skills can lead to an acceleration in the demand for skills (Acemoglu, 1998). So the timing of the increases in supply and demand is not a coincidence—instead, it reflects technology responding to the supply of skills. In this theory, rapid skill-biased technical change is not necessarily associated with rapid overall technical progress. In fact, an acceleration in skill bias could

[5] See, among others, Krueger (1993), Berman, Bound and Griliches (1994), and Autor, Katz and Krueger (1998) for evidence that the rapid spread of computers has increased the demand for skills. See Krusell, Ohanian, Rios-Rull and Violante (2000), Galor and Tsiddon (1997), Greenwood and Yorukoglu (1997), Aghion and Howitt (1998, chapter 9), Caselli (1999), Galor and Moav (2000), Violante (1999), Rubinstein and Tsiddon (1999), Aghion, Howitt and Violante (1999), and Gould, Moav and Weinberg (1999) for models in which rapid technical change increases the demand for skills and causes a rise in inequality.

[6] Naturally, supply and demand may have moved together because supply responded to demand. I argue below that the large increase in the supply of educated workers was not in anticipation, or in response to, high returns, but driven by a variety of other factors. More generally, I often focus on the effect of the supply of skills on technology not because I view supply as exogenous, but simply because the effect of supply on technology is more important in understanding the questions posed above. I discuss below how supply may respond to changes in skill premia, and how this may account for the joint behavior of the supply of, and demand for, skills over the past century.

cause a TFP slowdown because it creates an imbalance in the composition of R&D.

This approach also provides a possible explanation for the skill-replacing technical change of the nineteenth century. The emergence of the most skill-replacing technologies of the past two hundred years, the factory system, coincided with a large change in relative supplies. This time, there was a large migration of unskilled workers from villages and Ireland to English cities (see, for example, Habakkuk, 1962, Bairoch, 1988, or Williamson, 1990). This increase in the "*reserve army of unskilled workers*", slightly paraphrasing Marx, created profit opportunities for firms to exploit by introducing technologies that could be used with unskilled workers. In fact, contemporary historians considered the incentive to replace skilled artisans by unskilled laborers as a major objective of technological improvements of the period. Ure, a historian in the first half of the nineteenth century, describes these incentives as follows:

> "It is, in fact, the constant aim and tendency of every improvement in machinery to supersede human labor altogether, or to diminish its costs, by substituting the industry of women and children for that of men; of that of ordinary labourers, for trained artisans." (quoted in Habakkuk, 1962, p. 154).

These incentives for skill-replacing technologies, I argue, were shaped by the large increase in the supply of unskilled workers. So, it may be precisely the differential changes in the relative supply of skilled and unskilled workers that explain both the presence of skill-replacing technical change in the nineteenth century and skill-biased technical change during the twentieth century.

A major shortcoming of the "pure technological" approaches—of both exogenous and endogenous variety—is that they do not provide a natural explanation for the fall in the wages of low-skill workers. Although a variety of papers, including Caselli (1999), Greenwood and Yorukoglu (1997), and Galor and Moav (2000), show that technological revolutions may be associated with a fall in the wages of low-skill workers, it is difficult to see how sustained technological change can be associated with *an extended period* of falling wages of low-skill workers and stagnant average wages. This leads to the next question for this essay.

Why did the real wages of low-skill workers fall over the past several decades? There are a number of possible answers. First, labor market institutions, for example labor unions, have undergone important changes over the past 30 years, and these changes may have reduced the wages of many manufacturing workers, causing an increase in inequality and a decline in the real wages of low-skill workers (e.g., Freeman, 1991, DiNardo, Fortin and Lemieux, 1995, Lee, 1999). Second, international trade between skill-scarce less-developed countries and skill-abundant rich economies has increased over this period, and this may have put downward pressure on the wages of low-skill workers in the U.S. (e.g., Wood, 1994, Leamer, 1995). Third, there has been a transformation in the way firms are organized, or perhaps in the way that firms and workers match (see, for example, Acemoglu 1999a, Kremer and Maskin 1999, Bresnahan, 1997, Bresnahan et al, 1999, and Autor, Levy and Murnane 2000). I argue that neither one of these factors can by itself be the major cause of the recent changes in the wage structure. But technical change can also affect the organization of production and the institutions around it; for example, computers may lead to the emergence of organizations employing mostly skilled workers, with little need for unskilled workers, or technological developments may weaken the bargaining power of unions. In fact I believe that organizational change, labor market institutions and international trade have all interacted with technical change in a fundamental way; as a result, they amplified the direct effect of technical change on inequality, and likely caused the decline in the wages of less skilled workers.

Therefore, the overall picture that emerges is not necessarily one in which technology is the only factor affecting the distribution of income. On the contrary, the underlying thesis of this essay is that technology itself is no more than an endogenous actor. Instead, I argue that to explain the changes in the distribution of income, and to forecast what other changes may happen in the future, we need to understand the forces that shape technological progress, and how technology interacts with the overall organization of the labor market.

There is considerable uncertainty on many issues, and both more theoretical and empirical work is needed. Two areas deserve special attention. The first is the differential behavior of residual inequality and returns to schooling during the 1970s. Most economists view changes in residual inequality as related to changes in labor market prices. It is therefore puzzling that during the 1970s, while returns to schooling fell, residual and overall inequality increased. I argue below that models based on a single skill index (one type of skill or many types of skills that are perfect substitutes) are unable to explain this pattern. Instead, we need models with multi-dimensional skills. Moreover, for this type of models to explain the behavior of residual inequality during the 1970s and 1980s, technological progress needs to change the demand for different types of skills differentially. The endogenous technology models discussed above provide a possible explanation for why different dimensions of skills may have been affected differentially by technical change. Nevertheless, the reasons for this type of behavior require much more research. More generally, we know relatively little about the determinants of residual inequality, and this topic is a major research area for the future. The second area is cross-country differences in the behavior of wage inequality. While inequality increased sharply in the U.S., the UK and Canada, it increased much less in Germany and many Scandinavian economies. Although there are a number of recent papers addressing these questions, much uncertainty still remains. I conjecture that cross-country differences in wage inequality may reflect, in part, technological choices made by these countries in response to the different incentives created by their labor market institutions, but much more research on this topic is required.

2 Empirical Trends

The objective of this section is to illustrate a number of major inequality trends from the past several decades. My aim is not to offer a comprehensive survey of the empirical literature, but simply to put the most salient trends on the table to anchor the theoretical discussion (see, e.g., Gottschalk, 1997, Johnson, 1997, Katz and Autor, 2000, for recent surveys).

Figure 1 plots a measure of the supply of college skills between 1949 and 1995, constructed along the lines of Autor, Katz and Krueger (1998), as the ratio of college equivalents (those with college plus+0.5×those with some college) to noncollege equivalents (those with high school or less +0.5×those with some college).[7] It also plots returns to college. This picture summarizes many of the salient trends I want to emphasize.[8] In particular,

1. There has been a remarkable increase in the supply of skills in the U.S. economy over the past sixty years. In 1939, just over 6 percent of American workers were college graduates. By 1996 this number had increased to over 28 percent. In 1939, almost 68 percent of all workers did not have a high school degree. In 1996, this number had fallen to less than 10 percent (see, for example, Autor, Katz and Krueger, 1998, Table 1). The relative supply of skills plotted in Figure 1 provides a summary of these changes.

2. There has been no tendency for the returns to college to fall in the face of this large increase in supply—on the contrary, there is an increase in the college premium over this time period.

3. Following an acceleration in the supply of skills, returns to college fell sharply during the 1970s, leading Richard Freeman to conclude that "*Americans are over-educated*" (Freeman, 1976). Returns to college then rose very sharply during the 1980s. This increase in the returns to schooling has been one of the major motivating facts for the empirical inequality literature (e.g. Bound and Johnson, 1992, Katz and Murphy, 1992, etc.).

There have also been important changes in the overall distribution of wages. Figure 2 plots the 90th, 50th and 10th percentiles of the overall wage (weekly earnings) distribution for white male workers between 1963 and 1997 (with the 1963 values for all series indexed to 100).[9] This figure illustrates two more important patterns.

[7] See the Appendix for data details.

[8] An important issue is whether changes in wage inequality and returns reflect changes in the true returns to skills, or pure composition effects. In the Appendix, I show that pure composition effects cannot be responsible for these changes, and here I interpret them as changes in the true price of skills.

[9] Sample constructed as described in the Appendix. I focus here on wage inequality for white men since labor market participation of women increased substantially over the sample period, and this would likely

1. Overall wage inequality started to increase sharply in the early 1970s after a period of relative stability— prior to the 1970s, the 90th, 50th and 10th percentiles of the wage distribution followed each other closely, but came apart sharply in the 1970s.

2. Median wages stagnated from 1975 onwards, while workers at the 10th percentile of the wage distribution (i.e., "low-skill workers") saw their earnings fall in real terms to levels even below those in 1963.[10]

Figure 3 plots turns to another measure: residual (within-group) inequality, which shows inequality among observationally equivalent workers. This figure displays three measures of residual inequality among white male workers between 1963 and 1997: 50-10, 90-50 and 0.5 times 90-10 log wage residual differentials (I plot 0.5 times 90-10 wage differentials in order to fit this on the same scale as the other measures).

To calculate these measures, I look at the residuals from a standard Mincerian wage regression of the form

$$\ln w_{it} = X'_{it}\beta_t + v_{it}, \tag{1}$$

where w_{it} is weekly earnings for individual i observed in year t, and X_{it} is a set of controls which here include nine education dummies, a quartic in experience, and region controls (constructed from the March CPSs; see the Appendix for details of the sample). The fact that β_t is indexed by t indicates that returns to these observed characteristics are allowed to vary from year-to-year. The measures of residual inequality are calculated as the difference between the 90th and the 10th (or 50th and 10th, etc.) percentile values of the residual distribution from this regression, v_{it}. Residual inequality appears to have increased very much in tandem with overall inequality—it shows a sharp increase starting in the early 1970s.[11] Remarkably, all three measures of residual inequality behave very

contribute to the composition effects. Moreover, male-female wage difference narrowed substantially over the same time period as well. School quality for black men also underwent significant transformation (e.g., Welch, 1973, Card and Krueger, 1992), and this could create significant composition effects.

[10] Average wages, like median wages, have also stagnated. For example, white men aged 30-49 earned \$409 a week in 1999 dollars in 1949, and \$793 in 1969, which corresponds approximately to a 3.4 percent a year increase in real wages between 1949 and 1969. In contrast, the same age group earned \$909 in 1989, or experienced only a 0.6 percent a year increase between 1969 and 1989 (all numbers author's calculation from census data). The behavior of the median and average wage growth depends on the consumption deflator. I have followed the literature here in using the personal consumption expenditure deflator. It has been argued that this deflator overstates inflation because of difficulties in measuring quality change (e.g., Boskin, et al., 1995). Even in the presence of such measurement problems, unless there is an "acceleration" in this bias exactly around the 1970s, a large gap remains between the rate of increase of real wages before and after the 1970s.

Nevertheless, it has to be noted that part of this gap is due to the increase importance of nonwage income and benefits. In fact, thanks to the increase in benefits, the share of labor in national income has not fallen over this period (see, e.g., Krueger, 1999). So whether average wages have stagnated or continued to increase in line with output growth depends on how benefits are valued relative to earnings.

[11] DiNardo, Fortin and Lemieux (1995) show that wage inequality appears to increase starting in the

similarly, suggesting that forces affecting the top of the wage distribution (90-50) are also affecting the bottom of the wage distribution (50-10). Finally, note an important contrast between Figure 1 and Figures 2 and 3. While returns to schooling fell during the 1970s, overall and residual inequality increased. I return to this issue later in the essay.

3 Introduction to the Theory of Skill Premia

While, undoubtedly, many factors affect the distribution of wages, a natural starting point for an economic analysis is that of supply and demand. In the introduction to his pioneering study of income distribution, Tinbergen (1975, p. 15) wrote

> "...what matters is the difference between qualities *available* and qualities *required* by the demand side, that is by the organization of production." (italics in the original).

This is where I begin as well. I introduce a simple framework which links wages to supply of skills and to demand generated by the technology possibilities frontier of the economy. In this framework, there are two types of workers, skilled and unskilled (high and low education workers), who are imperfect substitutes. Imperfect substitution between the two types of workers is important in understanding how changes in relative supplies affect skill premia. For now I think of the unskilled workers as those with a high school diploma, and the skilled workers as those with a college degree. So the focus in this section is on returns to schooling (or between-group inequality), and I use the terms "skill" and education interchangeably. In practice, education and skills are only imperfectly correlated, so it is useful to bear in mind that since there are skilled and unskilled workers within the same education group, an increase in the returns to skills will also lead to an increase in within-group inequality.

Suppose that there are $L(t)$ unskilled (low education) workers and $H(t)$ skilled (high education) workers, supplying labor inelastically at time t. All workers are risk neutral, and maximize (the present value of) labor income. Also suppose that labor markets are competitive.[12]

The production function for the aggregate economy takes the form

$$Y(t) = F[K(t), L(t), H(t), t]$$

1980s in the May Current Population Survey (CPS) data. In the Appendix, I provide numbers from the survey by Katz and Autor (2000) showing consistent increases in wage inequality during the 1970s from March and May CPS data, and from census data.

[12] Although noncompetitive factors are likely to be important in accounting for wage differentials in European economies, competitive labor markets seem a good starting place in the analysis of wage inequality in the U.S. labor market.

where $K(t)$ is capital, and the production function explicitly depends on time to capture technical change. I have also imposed full employment. Although all the results of interest hold for a general constant returns to scale $F(.)$ function, to simplify the discussion I specialize it to the constant elasticity of substitution (CES) form,

$$Y(t) = \left[(A_l(t) L(t))^\rho + (A_h(t) H(t))^\rho\right]^{1/\rho}, \tag{2}$$

where $\rho \leq 1$. I also ignore capital. I drop the time argument when this causes no confusion.

The elasticity of substitution between skilled and unskilled workers in this production function is $\sigma \equiv 1/(1-\rho)$. I refer to skilled and unskilled workers as gross substitutes when the elasticity of substitution $\sigma > 1$ (or $\rho > 0$), and gross complements when $\sigma < 1$ (or $\rho < 0$). Three noteworthy special cases are: (i) $\sigma \to 0$ (or $\rho \to -\infty$) when skilled and unskilled workers will be Leontieff, and output can be produced only by using skilled and unskilled workers in fixed portions; (ii) $\sigma \to \infty$ when skilled and unskilled workers are perfect substitutes, and (iii) $\sigma \to 1$, when the production function tends to the Cobb Douglas case. The value of the elasticity of substitution will play a crucial role in the interpretation of the results that follow. In particular, in this framework, technologies either increase the productivity of skilled or unskilled workers, i.e., there are no explicitly skill-replacing or unskilled-labor-replacing technologies.[13] But, as we will see below, depending on the value of the elasticity of substitution, an increase in A_h can act either to complement or to "replace" skilled workers.

The production function (2) admits three different interpretations.

1. There is only one good, and skilled and unskilled workers are imperfect substitutes in the production of this good.

2. The production function (2) is also equivalent to an economy where consumers have utility function $[Y_l^\rho + Y_h^\rho]^{1/\rho}$ defined over two goods. Good Y_h is produced using only skilled workers, and Y_l is produced using only unskilled workers, with production functions $Y_h = A_h H$, and $Y_l = A_l L$.

3. A mixture of the above two whereby different sectors produce goods that are imperfect substitutes, and high and low education workers are employed in all sectors.

[13] A more general formulation would replace equation (2) with the production function

$$Y(t) = \left[(1-b_t)(A_l(t) L(t) + B_l(t))^\rho + b_t(A_h(t) H(t) + B_h(t))^\rho\right]^{1/\rho},$$

where B_l and B_h would be directly unskilled-labor and skill-replacing technologies, and an increase in b_t would correspond to some of the tasks previously performed by the unskilled being taken over by the skilled (see, e.g., Johnson and Stafford, 1999, on this). For most of the analysis here, there is little to be gained from this more general production function (but see Section 5.3).

Although the third interpretation is more realistic, I generally use one of the first two, as they are easier to discuss. Since labor markets are competitive, the unskilled wage is

$$w_L = \frac{\partial Y}{\partial L} = A_l^{\rho}\left[A_l^{\rho} + A_h^{\rho}(H/L)^{\rho}\right]^{(1-\rho)/\rho}. \tag{3}$$

This equation implies $\partial w_L/\partial H/L > 0$: as the fraction of skilled workers in the labor force increases, the wages of unskilled workers should increase. Similarly, the skilled wage is

$$w_H = \frac{\partial Y}{\partial H} = A_h^{\rho}\left[A_l^{\rho}(H/L)^{-\rho} + A_h^{\rho}\right]^{(1-\rho)/\rho},$$

which yields $\partial w_H/\partial H/L < 0$; everything else equal, as skilled workers become more abundant, their wages should fall.

Combining these two equations, the skill premium—the wage of skilled workers divided by the wage of unskilled workers—is[14]

$$\omega = \frac{w_H}{w_L} = \left(\frac{A_h}{A_l}\right)^{\rho}\left(\frac{H}{L}\right)^{-(1-\rho)} = \left(\frac{A_h}{A_l}\right)^{(\sigma-1)/\sigma}\left(\frac{H}{L}\right)^{-1/\sigma}. \tag{4}$$

Equation (4) can be rewritten in a more convenient form by taking logs,

$$\ln\omega = \frac{\sigma-1}{\sigma}\ln\left(\frac{A_h}{A_l}\right) - \frac{1}{\sigma}\ln\left(\frac{H}{L}\right). \tag{5}$$

Naturally, the skill premium increases when skilled workers become more scarce, i.e.,

$$\frac{\partial \ln\omega}{\partial \ln H/L} = -\frac{1}{\sigma} < 0. \tag{6}$$

This is the usual substitution effect, and shows that for *given skill bias of technology*, as captured by A_h/A_l, the relative demand curve for skill is downward sloping with elasticity $1/\sigma = (1-\rho)$. Intuitively, an increase in H/L can create two different types of substitutions. First, if skilled and unskilled workers are producing the same good, but performing different functions, an increase in the number of skilled workers will necessitate a substitution of skilled workers for tasks previously performed by the unskilled. Second, if skilled and unskilled workers are producing different goods, the greater number of skilled workers will lead to a substitution of the consumption of the unskilled good by the skilled good. In both cases, this substitution hurts the relative earnings of skilled workers.

Figure 4 draws the relative demand for skills as captured by equation (5) against the relative supply of skills, H/L, which is taken to be given for the purposes of this exercise.

[14]For some parameter values, skilled workers may have lower wages than the unskilled, i.e. $\omega \leq 1$. One may want to impose $\left(\frac{A_h}{A_l}\right)^{\sigma-1} > \frac{H}{L}$, to avoid this. Alternatively, one could assume that skilled workers can use the technologies normally used by the unskilled, A_l, and be more productive at this than the unskilled.

An increase in the relative supply, from H/L to H'/L', moves the equilibrium point along the downward sloping relative demand curve, and reduces the skill premium from ω to ω'.

An interesting case study of the response of the returns to schooling to an increase in the supply of skills is provided by the experience in the West Bank and Gaza Strip during the 1980s. As Angrist (1995) illustrates, there was a very large increase in the supply of skilled Palestinian labor as there opened Palestinian institutions of higher education, which were totally absent before 1972. Angrist shows that premia to college graduate workers (relative to high school graduates) that were as high as 40 percent quickly fell to less than 20 percent. The extent of substitution was also clear. First, many college graduate workers could not find employment in skilled jobs. Angrist (1995) shows a sharp increase in the unemployment rate of college graduates, and Schiff and Yaari (1989) report that only one in eight Palestinian graduates could find work in his profession, with the rest working as unskilled laborers, mainly in the construction industry. Second, premia for tasks usually performed by more educated workers fell sharply. Between 1984 and 1987, the premium for administrative and managerial jobs (relative to manual laborers) fell from .32 to .12, while the premium for clerical workers fell from .02 to -.08 (see Angrist, 1995, for details).

As equation (6) shows, the elasticity of substitution, σ, is important for the behavior of the skill premium when supply changes. The elasticity of substitution is also crucial for the response of the skill premium to changes in technology. Unfortunately, this parameter is rather difficult to estimate, since it refers to an elasticity of substitution that combines substitution both within and across industries. Nevertheless, there are a number of estimates using aggregate data that give a range of plausible values. The majority of these estimates are between $\sigma = 1$ and 2 (see, for example, Freeman, 1986).[15] The response of college premium for Palestinian labor reported in Angrist (1995), for example, implies an elasticity of substitution between workers with 16 years of schooling and those with less than 12 of schooling of approximately $\sigma = 2$.

Given the focus of this essay, it is useful to know how the skill premium responds to technology. Differentiation of (5) shows that the result depends on the elasticity of substitution. If $\sigma > 1$ (i.e., $\rho \in (0, 1]$), then

$$\frac{\partial \omega}{\partial A_h/A_l} > 0,$$

i.e., improvements in the skill-complementary technology increase the skill premium. This can be seen in Figure 4 as a shift out of the relative demand curve, which moves the skill premium from ω to ω''. The converse is obtained when $\sigma < 1$: that is, when $\sigma < 1$, an improvement in the productivity of skilled workers, A_h, relative to the productivity of

[15]These estimates are obtained from time-series or from cross-sectional data. The estimation strategies rely on a variety of assumptions, such as constant time trends in the demand for skills as in the work by Katz and Murphy (1992), so need to be interpreted with caution.

unskilled workers, A_l, shifts the relative demand curve in and reduces the skill premium. This case appears paradoxical at first, but is, in fact, quite intuitive. Consider, for example, a Leontieff (fixed proportions) production function. In this case, when A_h increases and skilled workers become more productive, the demand for unskilled workers, who are necessary to produce more output by working with the more productive skilled workers, increases by more than the demand for skilled workers. In some sense, in this case, the increase in A_h is creating an "excess supply" of skilled workers given the number of unskilled workers. This excess supply increases the unskilled wage relative to the skilled wage. This observation raises an important caveat. It is tempting to interpret improvements in technologies used by skilled workers, A_h, as "skill-biased". However, when the elasticity of substitution is less than 1, it will be advances in technologies used with unskilled workers, A_l, that increase the relative productivity and wages of skilled workers, and an increase in A_h relative to A_l will be "skill-replacing".

Nevertheless, the conventional wisdom is that the skill premium increases when skilled workers become relatively more—not relatively less—productive, which is consistent with $\sigma > 1$. In fact, as noted above, most estimates show an elasticity of substitution between skilled and unskilled workers greater than 1.

It is also useful to compute average wages in this economy. Without controlling for changes in the educational composition of the labor force, the average wage is

$$w = \frac{Lw_L + Hw_H}{L + H} = \frac{\left[(A_l L)^\rho + (A_h H)^\rho\right]^{1/\rho}}{1 + H/L}, \tag{7}$$

which is also increasing in H/L as long as the skill premium is positive (i.e., $\omega > 1$ or $A_h^\rho (H/L)^\rho - A_l^\rho > 0$). Intuitively, as the skill composition of the labor force improves, wages will increase.

Our results so far imply that in response to an increase in H/L:

1. Relative wages of skilled workers, the skill premium $\omega = w_H/w_L$, decreases.
2. Wages of unskilled workers increase.
3. Wages of skilled workers decrease.
4. Average wages (without controlling for education) rise.

These results can be easily generalized to the case in which physical capital also enters the production function, and the same comparative statics hold even when the economy has an upward sloping supply of capital. It is also useful to highlight the implications of an increase in A_h on wage levels. First, an increase in A_h, with A_l constant, corresponds to an increase in A_h/A_l; the implications of this change on the skill premium were discussed above. Moreover if A_h increases, everything else being equal, we expect both the wages of unskilled and skilled workers (and therefore average wages) to increase: technological improvements always increase all wages. This observation is important to bear in mind since, as shown in Section 2, the wages of low-skill workers fell over the past 30 years.

The most central result for our purposes is that as H/L increases, the skill premium, ω, should fall. In terms of Figure 4, the increase in supply corresponds to a rightward shift in the vertical line from H/L to H'/L', which would move the economy along the downward sloping demand curve for skills. But this tendency of the skill premium to fall could be counteracted by changes in technology, as captured by $\frac{\sigma-1}{\sigma}\ln(A_h/A_l)$. Therefore, this simple formulation encapsulates the essence of the two forces that Tinbergen (1975) emphasized;

> "The two preponderant forces at work are *technological development*, which made for a relative increase in demand and hence in the income ratio... and *increased access to schooling*, which made for a relative decrease", (p. 35, italics in the original).

As discussed in the empirical trends section, the past 60 years, and particularly the past 30 years, have witnessed a rapid increase in the supply of skills, H/L, but no corresponding fall in the skill premium. This implies that demand for skills *must have increased*—as a result of Tinbergen's "technological development"—to prevent the relative wages of skilled workers from declining. Although in richer models there could be other factors leading to such a steady increase in the demand for skills, the cause highlighted by this simple framework, skill-biased technical change, is the most natural candidate. More explicitly, the relative productivity of skilled workers, $(A_h/A_l)^{(\sigma-1)/\sigma}$, must have increased.

The increase in $(A_h/A_l)^{(\sigma-1)/\sigma}$ can be interpreted in a number of different ways. In a two-good economy, such skill-biased technical change corresponds to an increase in A_h/A_l and $\rho > 0$ ($\sigma > 1$)—i.e., skilled workers become more productive. Skill-biased technical change could also take the form of a decrease in A_h/A_l and $\rho < 0$ ($\sigma < 1$). In this case the "physical" productivity of unskilled workers would increase, but their relative wages would fall due to relative price effects. Alternatively, with the one-good interpretation, skill-biased technical change simply corresponds to an increase in $(A_h/A_l)^{(\sigma-1)/\sigma}$.

Some back-of-the-envelope calculations provide a sense of the rise in A_h/A_l implied by the changes in the structure of wages and employment. Autor, Katz and Krueger (1998) report employment and wage bill shares for different groups of workers in their Appendix Table A1. If we assume a specific value for σ, we can translate these numbers into changes in A_h/A_l. In particular, notice that the relative wage bill of skilled workers is given by

$$S_H = \frac{w_H H}{w_L L} = \left(\frac{A_h}{A_l}\right)^{(\sigma-1)/\sigma}\left(\frac{H}{L}\right)^{(\sigma-1)/\sigma}. \tag{8}$$

Hence, we have

$$\frac{A_h}{A_l} = \frac{S_H^{\sigma/(\sigma-1)}}{H/L}. \tag{9}$$

In Table 1, I calculate the implied A_h/A_l values for $\sigma = 1.4$ and for $\sigma = 2$ using workers with some college, college graduates, and college equivalents definitions of Autor, Katz and Krueger (1998)—see their paper for more a detailed analysis that controls for potential composition effects. In all cases, there is a very large implied increase in A_h/A_l and $(A_h/A_l)^{(\sigma-1)/\sigma}$. For example, the numbers indicate that, assuming an elasticity of substitution of 1.4, the relative productivity of college graduates, A_h/A_l, was approximately 0.030 in 1960, increased to 0.069 in 1970, and to 0.157 in 1980. Between 1980 and 1990, it increased by a factor of almost three to reach 0.470. As equation (5) shows, changes in the demand index $D = (A_h/A_l)^{\frac{\sigma-1}{\sigma}}$ may be more informative than changes in A_h/A_l, so Table 1 also gives the evolution of D.

The view that the post-war period is characterized by skill-biased technical change also receives support from the within-industry changes in employment patterns. With constant technology, an increase in the relative price of a factor should depress its usage in all sectors. Since the college premium increased after 1979, with constant technology, there should be fewer college graduates employed in all sectors—and the sectoral composition should adjust in order to clear the market. The evidence is very much the opposite. Berman, Bound and Griliches (1994) and Murphy and Welch (1993) show a steady increase in the share of college labor in all sectors.

This discussion leads to my first conclusion, which I highlight for future reference.

Conclusion 1 The past sixty years must have been characterized by skill-biased technical change.

Furthermore, Goldin and Katz (1998) provide evidence of technology-skill complementarity during the 1910s and 1920s. In light of this evidence, one might consider the bulk of the twentieth century to be characterized by skill-biased technical change, though whether technical change during the early twentieth century was skill-biased or not is not central for the focus of this paper.

4 Steady-Demand and Acceleration Hypotheses

The previous section highlighted the importance of skill-biased technical change over the past several decades. But why has technical change been skill-biased? And is there any pattern to the rate at which new technologies become more skill-biased? These are the questions I address in the rest of this essay. A first hypothesis is that the skill-biased technical change takes place steadily—at a constant pace—over time. Alfred Marshall begins *the Principles of Economics* by arguing that "Nature does not make jumps." It is then perhaps natural to begin with a hypothesis in which skill-biased technical change does not make jumps, but progresses steadily. The alternative would be a process which is at times more skill-biased than others—or even labor-biased during some episodes. In this section, I contrast the steady-demand hypothesis, which maintains that skill-biased technical change has progressed at a constant pace over the post-war period, against the acceleration hypothesis, which sees a break with past trends during recent decades.

4.1 Steady-Demand Hypothesis

According this hypothesis, there has been no major change in the structure of demand for skills. Versions of this story have been suggested by Freeman (1976), and it has been proposed as an explanation for the changes in the wage structure during the 1970s and the 1980s by Katz and Murphy (1992).

In a simple form, this hypothesis can be captured by writing

$$\ln\left(\frac{A_h(t)}{A_l(t)}\right) = \gamma_0 + \gamma_1 t, \tag{10}$$

where t is calendar time. Substituting this equation into (5), we obtain

$$\ln\omega = \frac{\sigma-1}{\sigma}\gamma_0 + \frac{\sigma-1}{\sigma}\gamma_1 t - \frac{1}{\sigma}\ln\left(\frac{H}{L}\right). \tag{11}$$

It is useful to link this equation to the two forces discussed above, and emphasized by Tinbergen (1975). According to equation (11), "technological developments" take place at a constant rate, but the supply of skilled workers could grow at different rates. Therefore, changes in the returns to skills are caused by uneven growth in the supply of skills. When H/L grows faster than the rate of skill-biased technical change, $(\sigma-1)\gamma_1$, the skill premium will fall, and when the supply growth falls short of this rate, the skill premium will increase. The story has obvious appeal since the 1970s, when returns to schooling fell sharply, were a period of faster than usual increase in the supply of college graduate workers as Figure 1 and Table 1 show. In contrast, the 1980s were a period of slow increase in the supply of skills relative to the 1970s. Katz and Murphy (1992) estimate a version of equation (11) above using aggregate data between 1963-1987.[16] They find

$$\begin{array}{lll} \ln\omega = & 0.033\cdot t & -0.71\cdot\ln\left(\frac{H}{L}\right) \\ & (0.01) & (0.15) \end{array}$$

This approach does fairly well in capturing the salient features of the changes in the college premium between 1963 and 1987.[17] In fact, Katz and Murphy show that the predicted

[16] They use the relative supply of college equivalent workers. This is defined as college graduates+0.29×some college-0.05×high school dropouts divided by high school graduates+0.69×some college+0.93×high school dropouts.

[17] More recently, Murphy, Riddle and Romer (1998) have argued this for Canada and the U.S., and Card and Lemieux (2000) for the U.S., Canada and the U.K..

values from the above equation are quite close to the observed movements in the college premium. This implies that we can think of the U.S. labor market since 1963 as characterized by an elasticity of substitution between college graduate workers and noncollege workers of about $\sigma = 1/0.71 \approx 1.4$, and an annual increase in the demand for skills at the rate of about 3.3 percent. The increase in the college premium during the 1980s is then explained by the slowdown in the rate of growth of supply of college graduates.

Nevertheless, there are a number of reasons for preferring a cautious interpretation of this regression evidence. The regression uses only 25 aggregate observations, and there is significant serial correlation in the college premium (as also noted by Katz and Murphy). If the true data were generated by an acceleration in skill bias and a larger value of the elasticity of substitution, this regression could estimate a smaller elasticity of substitution and no acceleration in the demand for skills (see below on this). For example, Katz and Murphy show that if the true elasticity of substitution is $\sigma = 4$, a significant acceleration in the skill bias of technical change is required to explain the data. Moreover, from the wage bill share data reported above, Autor, Katz and Krueger (1998) conclude that even for the range of the values for the elasticity of substitution between $\sigma = 1$ and 2, skill-biased technical change is likely to have been more rapid during the 1980s than the 1970s. This can also be seen in Table 4 above, where, for most measures, the increase in $(A_h/A_l)^{\frac{\sigma-1}{\sigma}}$ appears much larger between 1980 and 1990 than in other decades. I therefore do not consider this regression evidence as conclusive, and turn to discuss more detailed evidence on this issue.

4.2 Evidence on Steady-Demand versus Acceleration

The first piece of evidence often put forth in support of an acceleration relates to the role of computers in the labor market. Krueger (1993) has argued that computers have changed the structure of wages, and showed that workers using computers are paid more, and this computer wage premium has increased over time. Although this pattern is striking, it is not particularly informative about the presence or acceleration of skill-biased technical change. It is hard to know whether the computer wage premium is for computer skills, or whether it is even related to the widespread use of computers in the labor market. For example, DiNardo and Pischke (1997), and Enhorf and Kramartz (1998) show that the computer wage premium is likely to be a premium for unobserved skills.[18]

The second set of evidence comes from the cross-industry studies of, among others, Berman, Bound and Griliches (1994), Autor, Katz and Krueger (1998), and Machin and

[18] Equally, however, it would be wrong to interpret the findings of DiNardo and Pischke (1997) and Enhorf and Kramartz (1998) as evidence against and acceleration in skill-biased technical change, since, as argued below, such technical change would increase the market prices for a variety of skills, including unobserved skills.

Van Rennan (1998). These papers document that almost all industries began employing more educated workers during the 1970s and the 1980s. They also show that more computerized industries have experienced more rapid *skill upgrading*, i.e., they have increased their demand for college-educated workers more rapidly. For example, Autor, Katz and Krueger run regressions of changes in the college wage-bill share in three digit industries on computer use between 1984 and 1993. They find, for example, that

$$\Delta Sc_{80-90} = \underset{(.108)}{.287} + \underset{(.046)}{.147}\Delta cu_{84-93}$$

$$\Delta Sc_{90-96} = \underset{(.196)}{-.171} + \underset{(.081)}{.289}\Delta cu_{84-93}$$

where ΔSc denotes the annual change in the wage bill share of college graduates in that industry (between the indicated dates), and Δcu_{84-93} is the increase in the fraction of workers using computers in that industry between 1984 and 1993. These regressions are informative since the college wage bill share is related to the demand for skills as shown by equation (9). The results indicate that in an industry where computer use increases by 10 percent, the college wage bill share grows by about 0.015 percent faster every year between 1980 and 1990, and 0.03 percent faster in every year between 1990 and 1996.

Although this evidence is suggestive, it does not establish that there has been a change in the trend growth of skill-biased technology. As pointed out in Conclusion 1 above, the only way to make sense of post-war trends is to incorporate skill-biased technical change over the whole period. Moreover, Goldin and Katz (1998) present evidence suggesting that capital-skill complementarity may have been as high during 1910s as during the recent period because of increased demand for skills coming from the introduction of electricity in most manufacturing processes. Similarly, even though there were few computers in workplaces before the 1970s, other technological developments may have increased demand for skills as rapidly as —or more rapidly than—computers. Therefore, the question is whether computers and the associated information technology advances have increased the demand for skills *more* than other technologies did during the 1950s and 1960s, or even earlier. This question *cannot be answered* by documenting that computerized industries demand more skilled workers.

Cross-industry studies also may not reveal the true impact of computers on the demand for skills, since industries that are highly computerized may demand more skilled workers for other reasons as well.[19] In fact, when Autor Katz and Krueger (1998) run the above

[19] Doms, Dunne, and Troske (1997) show that new technologies (but not computers) are adopted by plants that have more skilled and more highly paid workers, and these plants do not increase their wages or demand for skills after the implementation of these technologies.

regressions for 1960-1970 college wage bill shares, they obtain

$$\Delta Sc_{60-70} = \underset{(.058)}{.085} + \underset{(.025)}{.071} \Delta cu_{84-93}$$

Therefore, industries investing more in computers during the 1980s were already experiencing more skill upgrading during the 1960s, before the arrival of computers (though perhaps slower, since the coefficient here is about half of that between 1980 and 1990). This suggests that at least part of the increase in the demand for skills coming from highly computerized industries may not be the direct effect of computers, but reflect an ongoing long-run shift towards more skilled workers. In this light, faster skill upgrading by highly computerized industries is not inconsistent with the steady-demand hypothesis.

The third, and probably most powerful, piece of evidence also comes from Autor, Katz and Krueger (1998). They document that the supply of skills grew faster between 1970 and 1995 than between 1940 and 1970—by 3.06 percent a year during the latter period compared to 2.36 percent a year during the earlier 30 years. In contrast, returns to college increased between 1970 and 1995 by about 0.39 percent a year, while they fell by about 0.11 percent a year during the earlier period. If demand for skills had increased at a steady pace, the skill premium should have also fallen since 1970.[20] Moreover, Autor, Katz and Krueger (1998) document greater within-industry skill upgrading in the 1970s, 1980s and 1990s than in 1960s, which is also consistent with more rapid skill-biased technical change during these later decades.

A simple regression analysis also confirms this point. I combined the data from the March CPSs and decennial censuses used in Figure 1 above. Using these data, a regression similar to that of Katz and Murphy for the period 1939-1996 yields similar results:

$$\ln \omega = \underset{(0.01)}{0.025 \cdot t} \quad \underset{(0.20)}{-0.56 \cdot \ln\left(\frac{H}{L}\right)},$$

with an R^2 of 0.63 and an implied elasticity of substitution of 1.8, which is somewhat larger than the estimate of Katz and Murphy. However, adding higher order terms in time (i.e., time squared, time cubed, etc.) improves the fit of the model considerably, and these higher-order terms are significant. In Figure 5, I plot the implied time trends from

[20] Returns to college fell between 1940 and 1970 because they are estimated to be very high in the 1940 census. There may be reasons to be suspicious of data quality from this census, because (i) the education variable was different, (ii) there may have been an overstatement of years of schooling, possibly by as much as a factor of 1.5 or 2 for some cohorts, and (iii) there was no self-employment income in this census. But it is not clear whether any of the measurement problems will cause an upward bias in the college premium. In any case, the level of the college premium from this census is not out of line with other historical evidence (see, e.g., Goldin and Katz, 2000). Moreover Autor, Katz and Krueger show that even ignoring data from the 1940 census, there is evidence for an acceleration in the skill bias of technical change. For example, for the range of the values for the elasticity of substitution between $\sigma = 1$ and 2, skill-biased technical change appears more rapid during the 1980s than in the 1970s and 1960s.

regressions with higher-order terms as well as the linear trend (all numbers were rescaled to fit in one graph). All three of these more flexible time trends show an acceleration in the relative demand for skills during the 1970s or 1980 (the quadratic and cubic time trends are almost identical, hence practically indistinguishable in the figure).

A fourth piece of evidence comes from Greenwood and Yorukoglu (1997) and Krusell, Ohanian, Rios-Rull and Violante (2000). These authors argue, based on the work of Griliches (1969), that equipment capital is more complementary to skilled workers than unskilled workers. This premise may be reasonable since advances in equipment often appear to substitute machines for tasks previously performed by unskilled workers. Following the work by Gordon (1990) and Greenwood, Hercowitz and Krusell (1997), these papers document that the post-war period has witnessed a secular decline in the relative price of equipment capital, and argue that the associated increase in the stock of equipment capital has led to skill-biased technical change. Moreover, they argue that this relative decline accelerated in the early 1970s, and the associated acceleration in the stock of equipment capital increased the demand for skills.

Krusell, Ohanian, Rios-Rull and Violante (2000) formalize their approach by assuming the following production function

$$Y = K_s^{\alpha}\left[b_1 L^{\mu} + (1-b_1)\left(b_2 K_e^{\lambda} + (1-b_2)\,H^{\lambda}\right)^{\mu/\lambda}\right]^{(1-\alpha)/\mu}$$

where K_s is structures capital (such as buildings), and K_e is equipment capital (such as machines). The parameter $\sigma_1 = 1/(1-\lambda)$ is the elasticity of substitution between equipment and skilled workers, and $\sigma_2 = 1/(1-\mu)$ is the elasticity of substitution between unskilled workers and the equipment-skilled worker aggregate. If $\sigma_1 > \sigma_2$ (i.e., $\mu > \lambda$), equipment capital is more complementary to skilled workers than unskilled workers, and as a result, an increase in K_e will increase the wages of skilled workers more than the wages of unskilled workers. More formally, the skill premium in this model is

$$\omega = \frac{w_H}{w_L} = \frac{(1-b_2)(1-b_1)\,H^{\lambda-1}\left(b_2 K_e^{\lambda} + (1-b_2)\,H^{\lambda}\right)^{(\mu-\lambda)/\lambda}}{b_1 L^{\mu-1}}. \tag{12}$$

Differentiation of (12) shows that as long as $\mu > \lambda$, $\partial\omega/\partial K_e > 0$. So provided that equipment capital is more complementary to skilled workers than unskilled workers, an increase in the quantity of equipment capital will increase the demand for skills. Since the post-war period has been characterized by a decline in the relative price of equipment goods, there will be an associated increase in the quantity of equipment capital, K_e, increasing the demand for skills steadily.

Figure 6, which plots the log of this relative price series, shows the faster proportional decline after the 1970s. The behavior of the relative price series then suggests that there may have been an acceleration in the substitution of equipment capital for labor, causing more rapid skill-biased technical change.

Nevertheless, there are serious difficulties in adjusting capital prices for quality. This suggests that we may want to be cautious in interpreting this evidence. Another problem comes from the fact that, as I will discuss in more detail below, a variety of other evidence does not support the notion of faster technological progress since 1974, which is a basic tenet of this approach. Finally, one would presume that if, in fact, the decline in the relative price of equipment capital is related to the increase in the demand for skills, then in a regression of equation (11) as in the work by Katz and Murphy (1992), it should proxy for the demand for skills and perform better than a linear time trend. Table 2 reports a series of regressions which show that, on the contrary, the level or the log of the relative price of equipment capital is not significant in such regressions. Column 1 shows the equivalent of the regression by Katz and Murphy (1992) with only a time trend and the relative supply of skills. Columns 2 and 3 show regressions that replace the time trend with the level and log of the relative price of equipment capital. These terms are significant, but the fit of the regression is worse than the one in Column 1. The remainder of the table shows that once these terms are entered simultaneously with a time trend, the time trend is significant, while there is no evidence that the relative price of equipment capital matters for the demand for skills. This evidence casts some doubt on the view that the relative price of equipment capital is directly linked to the demand for skills and that its faster decline since 1970s indicates an acceleration in skill bias.

A final piece of evidence comes from the behavior of overall and residual inequality over the past several decades. As Section 2 documented, overall wage inequality was fairly stable until the 1970s. Since 1970, both overall and residual wage inequality have risen sharply. This increase in inequality weighs in favor of a marked change in labor market prices and demand for skills. This argument is based on the interpretation of changes in residual inequality as reflecting changes in labor market prices, a thesis put forth by Juhn, Murphy and Pierce (1993).[21] A concrete example might clarify why residual inequality is linked to the demand for skills. Suppose that two otherwise identical individuals differ in terms of their unobserved skills (for example, in terms of interpersonal skills, motivation, specific skills for their job, or IQ).[22] Denote the unobserved skill of individual 1 by a_1 and

[21] Of course, and alternative—and more cynical— view would be to interpret residual inequality as "a measure of our ignorance". When a standard wage regression such as (1) provides a good fit, the residuals will be less disperse. Nevertheless, given the variety of skills that we are unable to measure in standard data sets, much of the residual will plausibly reflect rewards to some unobserved skills.

[22] By unobserved skills, I mean skills that are not observed by the econometrician. These skills could be—and are likely to be— observed by employers. This type of skills are often referred to as "unobserved ability". This does not imply that these unobserved skills are necessarily synonymous with IQ or other single dimensional skill indices.

that of individual 2 by $a_2 > a_1$, and assume that wages are given by

$$\ln w_{it} = 2\theta_t a_i + \gamma_t h_i, \tag{13}$$

where γ_t is the price of h skills at time t, while θ_t is the price of a skills. Since these individuals are identical in all respects other than their unobserved skills, a, we have that the variance of log wages (or of residual wages) among these two individuals is

$$Var(\ln w) = \theta_t^2 (a_2 - a_1)^2 .$$

Now if at a later date, t', this variance increases to $Var(\ln w)'$, and we know that these two individuals are still identical in all other respects and that $a_2 - a_1$ has not changed, we can interpret the increase in $Var(\ln w)$ as reflecting an increase in the price of unobserved skills, θ_t. The discussion in the Appendix shows that the bulk of the increase in overall and residual inequality cannot be explained by composition effects, so this increase is most likely due to a rise in the price of and demand for unobserved skills during the 1970s.

Overall, there is a variety of evidence suggesting an acceleration in skill bias over the past 25 or 30 years. Although not all evidence is equally convincing, the rise in the returns to schooling over the past 30 years, despite the very rapid increase in the supply of skills, and the behavior of overall and residual inequality since the 1970s suggest a marked shift in the demand for skills over the past several decades. I therefore tentatively conclude:[23]

Conclusion 2 The behavior of returns schooling and residual inequality over the past three decades suggest an acceleration in the demand for skills beginning in the 1970s or 1980s.

5 Acceleration in Skill Bias

What explains the more rapid increase in the demand for skills over the past several decades? There are a number of alternatives. The first is a change in labor market institutions. The second is the role of increased international trade. The third is more rapid skill-biased technical change. I argue in Section 6 that changes in labor market institutions and the increased importance of international trade cannot explain the change in labor market prices by themselves. Moreover, the evidence discussed in Section 4 is consistent with new technologies playing an important role in changing the wage structure. So here I begin with changes in technologies, and in particular discuss "pure technological" approaches where technology is the only factor determining the demand for skills.

[23] It should be noted that the pace of skill-biased technical change may have been slower during the 1990s (see Autor, Katz and Krueger, 1998).

5.1 "Technological Revolutions" and Acceleration in Skill Bias

The first group of technological theories link the acceleration in skill bias to exogenous technological developments, and argue that a "technological revolution" led to more rapid skill-biased technical change beginning in the 1970s or 1980s. In terms of the model developed above, this corresponds to a more rapid increase in A_h/A_l during this period, translating into greater skill premia. Many of the proponents of this view argue that the acceleration in skill bias is, at least in part, related to information technology and computers (for example, Krueger, 1993, Berman, Bound and Griliches, 1994, Autor, Katz and Krueger, 1998, Berman, Bound and Machin, 1998).

An interesting version of this story is the one developed by Krusell, Ohanian, Rios-Rull and Violante (2000). As pointed above, these authors argue, based on the work of Gordon (1990) and Greenwood, Hercowitz and Krusell (1997), that the demand for skills accelerated as a result of the more rapid decline in the relative price of capital equipment. They argue that this relative decline accelerated beginning in the early 1970s.[24] The Krusell et al theory is attractive since it provides a unified framework in which we can identify both the cause of the steady increase in the demand for skills, and the source of the more rapid skill-biased technical change, though the evidence provided in the previous section casts some doubt on the link between the relative price of equipment and demand for skills.

The main idea of these approaches is that new technologies are more complementary to skilled workers than to unskilled workers—for example, there are more rapid advances in the technologies used by skilled workers, as captured by A_h above. Rapid technological progress then corresponds to an *acceleration in skill bias*. An alternative perspective, building on an idea originally suggested by Nelson and Phelps (1966), emphasizes the ability of skilled workers to deal with the introduction of new technologies. According to this view, demand for skills will automatically increase during periods of rapid technological change. Welch (1970) gave an early succinct summary of these two views. The first view—the acceleration hypothesis—would maintain that

> "technical change may not be neutral between skill classes. It may be that increments in technology result in increments in the relative productivity

[24] I classify this approach as one of exogenous technology, since the driving force, the decline in the relative price of equipment capital, is assumed exogenous.

An alternative interpretation of the approach by Krusell Ohanian, Rios-Rull and Violante (2000) is that the main determinant of the demand for skills is not technology-skill complementarity, but capital-skill complementarity. I believe that the distinction between technology-skill versus capital-skill is not very useful. Capital-skill complementarity could play an important role only in a model as in Greenwood et al (1997) or Krusell et al (2000), where new capital embeds superior technologies. In this sense, it would be a combination of new capital and new technologies that is increasing the demand for skills. Moreover, Autor, Katz and Krueger (1998) show that demand for more educated workers across industries is affected by high-tech capital (e.g., computers), but not by equipment capital, suggesting further that it is new technologies, not simply capital-intensity, that matters for inequality.

of labor that are positively related to skill level." (p. 38).

In contrast, the second view—the Nelson-Phelps hypothesis—argues that

> "...the productivity of education would be positively related to the rate of change in useful technology (the ability to change) and to the size of the technological gap (room for innovation). In this case, if the rate of utilization of technology is accelerating, or if the technology gap is growing, the return to education will rise relative to other inputs." (p. 38).

Studies building on the Nelson-Phelps insight include Galor and Tsiddon (1997), Greenwood and Yorukoglu (1997), Caselli (1999), Aghion and Howitt (1998, chapter 9), Galor and Moav (2000), Violante (1999), Rubinstein and Tsiddon (1999), Aghion, Howitt and Violante (1999), and Gould, Moav and Weinberg (1999).[25] These papers argue that there has been a technological revolution in the U.S. economy starting in the 1970s, and relate the rise in inequality to the increased demand for skills resulting from the technological revolution. For example, Greenwood and Yorukoglu (1997) draw a parallel between the first and the second industrial revolutions and what has been happening in the U.S. economy since 1974. Caselli (1999) develops a similar theory where a technological revolution increases the demand for workers who can switch to the sectors that benefit from the introduction of new technologies.[26]

To get a basic understanding of these approaches, it is useful to consider a simplified version of the model by Galor and Moav (2000) adapted to the above framework. Suppose that

$$A_l = \phi_l(g)a \text{ and } A_h = \phi_h a \tag{14}$$

where a is a measure of aggregate technology, and g is the growth rate of a, i.e., $g \equiv \dot{a}/a$. The presumption that skilled workers are better equipped to deal with technological progress can be captured by assuming that $\phi_l' < 0$. Galor and Moav (2000) refer to this assumption as the "erosion effect," since it implies that technical change erodes some of the established expertise of unskilled workers, and causes them to benefit less from technological advances than skilled workers do. Substituting from (14) into (4), the skill premium is

$$\omega = \frac{w_H}{w_L} = \left(\frac{A_h}{A_l}\right)^{(\sigma-1)/\sigma} \left(\frac{H}{L}\right)^{-1/\sigma} = \left(\frac{\phi_h}{\phi_l(g)}\right)^{(\sigma-1)/\sigma} \left(\frac{H}{L}\right)^{-1/\sigma}. \tag{15}$$

[25] See Aghion (2000) for an approach that combines the Nelson-Phelps insights with the Schumpeterian notion of creative-destruction to discuss the impact of the diffusion of computers on inequality.

[26] The explanation offered by Violante (1999), Rubinstein and Tsiddon (1999), and Aghion, Howitt and Violante (2000) is somewhat different. They argue that there is increased uncertainty at times of rapid technological change, and more skilled workers are better able to cope with uncertainty. This idea is also related to a thesis first put forth by Piore and Sabel (1984) that the oil price shocks increased the uncertainty faced by producers, and induced them to change the organization of production.

Therefore, as long as $\phi_l' < 0$, more rapid technological progress, as captured by a higher level of g, will increase the skill premium.[27]

Theories that explain the increase in inequality as a result of rapid technological progress have a number of attractive features.[28] First, many economists and commentators view the advances in computer and information technology as a break with the technologies of the past, and so are open to the idea that we might be in the midst of a technological revolution. Second, a variety of evidence supports the notion that skilled workers have a comparative advantage in coping with rapid technical change. Bartel and Lichtenberg (1987) show that firms introducing new technologies hire more skilled workers. Bartel and Sicherman (1998) document that returns to unobserved ability appear to be higher in industries with more rapid technical change. Foster and Rosenzweig (1996) provide evidence from developing countries that more educated workers are better placed to take advantage of advances in agricultural technology.

The main difficulty with both the theories based on the acceleration and the Nelson-Phelps hypotheses is that they rely very explicitly on rapid technical change in recent decades. There is little direct evidence that the decades between 1970 and 1995 have been a period of rapid technical change. First, this period has experienced sluggish TFP and output growth relative to earlier periods. Greenwood and Yorukoglu (1997) and Hornstein and Krusell (1996) argue that the slow TFP growth itself may be an outcome of the more rapid technical change. According to this argument, new revolutionary technologies first reduce productivity growth as firms and workers spend their time learning to use these technologies. Moreover, following a suggestion by Griliches, many have argued that our ability to measure TFP growth may have deteriorated following a change in technological regime. Neither of these arguments are very convincing, however.

It is difficult to imagine how a new and radically more profitable technology will first lead to twenty five years of substantially slower growth. Although, in an influential paper, Paul David (1990) argues that the spread of electricity to American manufacturing was also slow and productivity gains from electrification were limited until the 1920s, the parallel with the recent productivity slowdown should not be overstated. First, though productivity growth from electrification was sluggish during the early 1900s, the U.S. economy overall had a much higher level of output growth than growth levels experienced over the past three decades. Data from Table A-XVIII of Kendricks (1961) imply that output growth

[27] Galor and Moav (2000) discusse in detail the response of workers of different ability to technological change, which I ignore, since this issue is not essential to the argument here.

[28] These theories also predict that inequality should increase when new technologies are being introduced, but should decline when these new technologies are standardized and being used routinely by many firms (see, for example, Galor and Tsiddon, 1997, or Aghion, Howitt and Violante, 2000). So far, there seems to be no evidence of a decline in inequality in the U.S., but perhaps the years to come will see a return to the levels of inequality experienced during the 1960s, vindicating this approach.

between 1899 and 1909 in the U.S. economy was 4.2 percent a year, while between 1909 and 1919, it was 3 percent, and between 1990 and 1929, output grew by 3.6 percent a year. Second, as noted by Oliner and Sichel (1994), computers and other advanced office equipment have only been a trivial part of the aggregate capital stock of the U.S. economy until the mid 1990s. It is therefore unlikely that the whole of the U.S. economy has been adapting to the changes in this relatively small part of the capital stock Finally, as shown by Brendt, Morrison and Rosenblum (1994), more computerized sectors did not perform any better in terms of labor productivity growth over this period, and this pattern is also difficult to reconcile with a computer-led technological revolution.

It is also useful to note that although computers have no doubt increased our standards of living and quality of life over the past thirty years, it seems (at least to me) that they are much less radical innovations than certain previous new technologies. To gain perspective, consider the difference that the telegraph makes to a world in which the fastest medium of communication were pigeons. Mokyr (1990, p. 124) describes this as follows:

> "The telegraph had an enormous impact on 19th-century society—possibly as great as that of the railroad. Its community and political value was vast, as was its effect in coordinating international financial and commodity markets. Unlike the railroad, it had no close substitutes, the closest being homing pigeons and semaphore."

Or consider the difference that the automobile and air conditioning made to the quality of life, and electricity and interchangeable parts made to the manufacturing sector. As also pointed out by Gordon (1998), compared to these improvements, the switch from mainframes to PCs, or from telephone to e-mail, or from the typewriter to the word processor seem more modest.

The argument that we have become worse at measuring productivity growth is also not totally compelling. As Bailey and Gordon (1992) document, productivity slowdown has been concurrent in many sectors, some of them with little problems in measuring output or output quality. It is interesting to note, however, that evidence in favor of this hypothesis may yet emerge, especially since productivity growth has been quite rapid during the past three years (but see Gordon 1998, and more recently, Jorgensen and Stiroh, 2000, on this).

In any case, historical evidence is not necessarily in line with a view that times of rapid technical change increase inequality. As discussed in the introduction, the major technological changes of the nineteenth century appear to have been largely unskilled-biased and to have reduced inequality, even though they seem as radical as computer technology. This suggests that it is the skill bias of technology, not merely its rapid arrival, that is important for the demand for skills.

A final problem for all of the approaches based exogenous technological developments is the coincidence in the timing of this change, and the rapid increase in the supply of skilled workers. Recall that there was a very large increase in the supply of college graduate workers during the late 1960s and early 1970s (Figure 1 and Table 1 show the large increase in the employment share of college workers between 1970 and 1980). So the acceleration in skill bias is either concurrent with, or immediately follows, this large increase in the supply of skills. There is no a priori reason to expect the acceleration in skill bias to coincide with the rapid increase in the supply of skills. Those who want to subscribe to the exogenous technological progress view have to explain this as a chance event.[29]

5.2 Endogenous Skill-Biased Technical Change

The theories discussed so far presume technical change to be skill-biased *by nature* (or, at the very least, recent technical change to have been skill-biased). A different perspective is to link the type of technologies that are developed and adopted to (profit) incentives, or to *demand pull* as emphasized by Schmookler (1966). This is also the approach taken by the endogenous growth theory, which determines the overall rate of technical change—but not the degree of skill bias—from profit incentives (e.g., Romer, 1990, Grossman and Helpman, 1991, Aghion and Howitt, 1992).

Historical evidence is consistent with the notion that profit incentives and opportunities are important for the development and introduction of new technologies. Braudel (1984, p. 566) took a strong position on this:

> "the efficient application of technology lags, by definition, behind the general movement of the economy; it has to be called on, sometimes several times, to meet a precise and persistent demand."

An interesting example of the timing of technological development responding to profit incentives is given by the introduction of the electric street car in U.S. cities during the late nineteenth century. In his history of electricity in the U.S., Nye (1990) describes this as follows: "Cities grew larger, better transportation was needed, so the [electric] trolley was invented, called into being by the crowded late nineteenth century cities....By the 1870s large cities had ceased to be accessible by foot, or built to the scale of pedestrians, and traffic congestion was terrible." (p. 85). This created the profit opportunities to develop

[29] One could argue that the supply of skilled workers increased because, during the 1960s, workers anticipated that there was going to be a technological discontinuity in the decades to come, and responded to this by increasing their education. This story appears quite unreasonable, however. There is no evidence that anyone, let alone teenagers, foresaw the technological developments of the 1970s and the 1980s as early as the 1960s. Moreover, the increase in the supply of skills can be explained by two factors. First, the Vietnam era draft laws encouraged young males to stay in college longer (and indirectly also influence female enrollments). Second, college enrollments were on an upward trend since the early 1950s, and much of the increased supply of college graduate workers is accounted for by the interaction of this upward trend and the very large relative size of the baby boom cohorts.

and introduce the electric trolley. The technological requirements had been met long before, and awaited these profit opportunities. Nye writes "However great the need for the electric trolley after 1870, it was hardly a new idea; it had been the object of experiment during four decades." (p. 86).

Another example of the type of innovation responding to profit incentives is provided by the cotton gin. In the late eighteenth and early nineteenth centuries, Britain imported most of its cotton from the West Indies, Brazil and India, whereas only green seed cotton, which was more difficult to clean, could be grown in most of the American South. A machine to remove the seeds was essential for the success of American cotton. In contrast to almost all other textile innovations that were taking place in England and Europe, such a machine, the cotton gin, was developed in the U.S. in 1793 by Eli Whitney in response to this need. The impact of the cotton gin on the South was nothing short of spectacular. In the court case over the patent rights, Judge Johnson wrote:

> "[...as a result of the cotton gin]... individuals who were depressed with poverty, and sunk with idleness, have suddenly risen to wealth and respectability. Our debts have been paid off, our capital increased; and our lands are treble in value." (Quoted in Green, 1956, p. 92).

Within a short time, Eli Whitney's gin turned the U.S. from a cotton importer into the largest cotton exporter.

Schmookler (1966) provides a famous argument for the importance of *demand pull* in the development of many technologies. He documents rapid innovations in railroads following increased purchases of railroad equipments, and more generally argues that industries with greater investments experience faster technological progress because the returns to such progress are greater. A natural next step is then to argue that the degree of skill bias in technical change is also determined by profit opportunities and by the demand for different types of technologies. Here, by endogenous (skill bias) technology approach I mean the view that the degree of skill bias in technical change is influenced by profit incentives.

A key determinant of profitability is market size. As Schmookler (1966) stated in the title of two of his chapters: *"The amount of invention is governed by the extent of the market."* The most successful businessmen have always been aware of this. For example, Matthew Boulton wrote to his business partner, James Watt, "It is not worth my while to manufacture your engine for three countries only, but I find it very well worth my while to make it for all the world" (quoted in Scherer, 1984, p. 13). Schmookler (1966) similarly provided many examples where market size was crucial in determining the directions of technical change. The horseshoe is perhaps the most interesting one. Schmookler documented that there was a very high rate of innovation throughout the late nineteenth and early twentieth centuries in this very ancient technology, invented in the second century B.C., and no tendency for inventors to run out of additional improvements. On the contrary, inventions and patents increased because demand for horseshoes was high. Innovations came to an end only when "the steam traction engine and, later, internal combustion engine began to displace the horse..." (p. 93).

According to this reasoning, the development of skill-biased technologies will be more profitable when they have a larger market size—i.e., when there are more skilled workers. Therefore, the equilibrium degree of skill bias could be an increasing function of the relative supply of skilled workers. An increase in the supply of skills will then lead to skill-biased technical change. Furthermore, an acceleration in the supply of skills can lead to an acceleration in the demand for skills. It is useful to link this approach to technological development to the above framework: this framework explained the prices of skills by supply and technology, while the perspective of endogenous skill bias relates technology to the supply of skills.[30] Tinbergen in his pioneering study of this supply-demand framework, in fact, foresaw this possibility, and wrote (1975, p. 61): "...an inequality-furthering phenomenon is technological development. But need it be? Increasingly we get the feeling that technological development is not simply something given, but that it may be guided, within limits."

At some level, the idea that there will be more technologies developed, created and adopted for skilled workers—"within limits"— when there are more skilled workers around is quite appealing. An extreme form of this view would be that captured by my model in Acemoglu (1998). There, forward looking profit maximizing firms create new technologies anticipating the profitability of these different investments. According to this view, it would be the Vietnam War draft laws and the high college enrollment rates of the baby boom cohorts that *induced* the development of computers. Obviously such an interpretation is not literal. A more plausible interpretation is that new technological platforms—*macroinventions* to use Joel Mokyr's term or *General Purpose Technologies* to use Bresnahan and Trajtenberg's term —stem from advances in basic science or from labs with little profit maximizing incentives. The development of the microchip would be such a macroinvention. But what matters for most workers in the labor force is how this new technological platform is developed, i.e. the *microinventions* that follow the macroinvention. At the expense of oversimplifying, we can say that the microchip could have been used to develop advanced scanners that would increase the productivity of unskilled workers, or advanced computer-assisted machines that would be used by skilled workers to replace unskilled workers. The theory of endogenous skill bias requires that the extent of the advances in these two technologies are affected by profit opportunities. When there are more college graduates, computers become relatively more profitable to develop than scanners,

[30] Naturally, it is also possible to link the supply of skills to skill premia. See below for a discussion.

and this explains the acceleration in skill bias.[31]

The endogenous response of firms to the increase in supply will raise the demand for skills. In fact, supply may not simply create its own demand, but the response of firms could be so pronounced that demand could *overshoot* the supply. In this theory, therefore, the increased supply may be the cause of the increased skill premia (see Acemoglu, 1998, and also Kiley, 1999). Here I outline a simplified version of this theory based on the above framework.

Suppose that consumers have a utility function defined over $Y = [Y_l^{\rho} + Y_h^{\rho}]^{1/\rho}$, and that $Y_h = N_h H$ and $Y_l = N_l L$ where N_h and N_l can be interpreted as the number of specialized machines used with skilled and unskilled workers, respectively. This is equivalent to the above setup with $A_h = N_h$ and $A_l = N_l$. An increase in N_h relative to N_l will correspond to skill-biased technical change as long as $\sigma = 1/(1-\rho) > 1$. From consumer maximization, the relative price of skill-intensive goods is

$$p \equiv \frac{p_h}{p_l} = \left[\frac{N_h H}{N_l L}\right]^{\rho-1}, \tag{16}$$

where once again p_h denotes the price of good Y_h and p_l is the price of Y_l.

Suppose now that these specialized machines are created and sold by profit maximizing monopolists. Creating a new machine costs B units of the final good Y, and the marginal cost of producing these machines, once created, is zero. The marginal willingness to pay for an additional machine in the two sectors are given by the derivatives of $p_h Y_h$ and $p_l Y_l$ with respect to N_h and N_l, i.e.,

$$p_h H \text{ and } p_l L. \tag{17}$$

I assume that the creator of each new machine obtains this "market" marginal willingness to pay. Equation (17) therefore highlights two effects that encourage the creation of new technologies.

1. The price effect: technologies producing more expensive goods will be improved faster. Since goods using the scarce factor will command a higher price (see equation (16)), this effect implies that there will be more innovation directed at the scarce factor—i.e., directed at unskilled workers during the 1970s and 1980s.

[31] There is in fact some evidence that the composition of R&D has shifted towards more skill-biased technologies during the period of the rapid increase in the supply of college-educated workers. From the R&D expenditure data reported by the NSF, in 1960 company funded R&D for office computing was 3 percent of the total company funded R&D expenditure. This ratio has increased to 13 percent by 1987, suggesting that during this period of rapid increase in the supply of skills, there has been significantly more R&D directed to one of the technologies most complementary to skills.

2. The market size effect: a larger clientele for a technology leads to more innovation. Since the clientele for a technology is effectively the workers who use it, the market size effect encourages innovation for the more abundant factor, and encourages more technologies for skilled and highly educated workers during the 1970s and 1980s.

The creation of new machines will stop when the marginal increase in profits is equal to the marginal cost of innovation in both sectors. This implies that in equilibrium

$$\frac{p_h H}{p_l L} = 1, \tag{18}$$

i.e., the price and market size effects have to be balanced in equilibrium. How can equation (18) be satisfied? Since H/L is fixed, equation (18) can only hold if the relative price of skill-intensive goods, $p = p_h/p_l$, adjusts. From equation (16), this can only happen if N_h/N_l change. Therefore, in this economy, the skill bias of technology has to adjust in order to *"clear the technology market"*. Combining (18) and (16), we obtain equilibrium skill bias as

$$\frac{N_h}{N_l} = \frac{A_h}{A_l} = \left(\frac{H}{L}\right)^{\rho/(1-\rho)}. \tag{19}$$

This equation shows that when $\rho > 0$, i.e., when skilled and unskilled good are gross substitutes, the market size effect will dominate the price effect, and a greater relative supply of skilled workers will lead to more skill-biased technologies— higher N_h/N_l.

Finally, the skill premium in this economy is given by

$$\omega = \frac{p_h N_h}{p_l N_l} = \left(\frac{H}{L}\right)^{(2\rho-1)/(1-\rho)}$$

where the final expression is obtained by substituting from equation (19).

The most important result is that if $\rho > 1/2$, i.e., if the elasticity of substitution σ is greater than 2, the skill premium will be an *increasing* function of the relative supply of skills.[32] This is because an increase in H/L encourages so much skill-biased technical change that the demand for skills increases more than enough to offset the increase in the supply of skills. As a result, the (long-run) relative demand for skills is an upward sloping curve as drawn in Figure 7, and an increase in the supply of skilled workers *increases* the skill premium.

There are a number of implications that follow from this approach. First, as the relative supply of skilled workers has been growing throughout the past sixty years, we expect

[32] The result that the elasticity of substitution needs to be greater than 2 for the long-run relative demand to slope upwards is a feature of the simple model here, and does not generalize to richer environments. In any case, there are a number of estimates above 2, and a somewhat upward sloping relative demand curve for skills is an empirical possibility.

technology to endogenously respond by becoming more skill-biased over time. If the elasticity of substitution between skilled and unskilled workers is greater than 2, i.e., $\rho > 1/2$, the increase in the demand for skilled workers would be more than enough to offset the increase in the supply of skilled workers, and the economy would be moving steadily along an upward sloping relative demand curve for skills. This would explain why returns to college have been increasing over the past half century.

A new theory for the acceleration in skill bias also emerges from this simple model. According to this theory, the rapid increase in the supply of college educated workers during the 1970s created a more pronounced shift towards skill-biased technologies, increased the demand for skills further, and raised the college premium. This story becomes more interesting once we recognize that the equilibrium skill bias of technologies, N_h/N_l, is a sluggish variable determined by the slow buildup and development of new technologies. In this case, a rapid increase in the supply of skills would first reduce the skill premium as the economy would be moving along a constant technology (constant N_h/N_l) curve as drawn in Figure 7. After a while the technology would start adjusting, and the economy would move back to the upward sloping relative demand curve, with a very sharp increase in the college premium. This theory therefore gives an interpretation for both the decline in the college premium during the 1970s and the subsequent large surge, and relates both to the large increase in the supply of skilled workers.

For the key insights of this theory, that increases in the relative supply of skills induce skill-biased technical change, we *do not need* the long-run relative demand curve to be upward sloping. When $\rho < 1/2$, increases in the supply of skills still induce skill-biased change, but this technical change would not be enough to prevent the skill premium from falling. Further "exogenous" skill-biased technical change is also necessary to explain why returns to schooling have risen over the past 60 years. With a downward sloping long-run demand curve, the story for the 1970s and 1980s would be different as well. The large increase in the supply of skills again moves the economy along a steeply downward sloping constant technology demand curve. The response of technology then moves the economy to a less steep long-run demand curve as drawn in Figure 8, raising the skill premium.

There are also other historical episodes in which a large increase in the supply of skills appears to have affected the direction of technical change. High school enrollment and graduation rates doubled in the 1910s. Goldin and Katz (1995) argue that increased enrollments were mostly driven by supply side factors; changes in the location and curricula of schools and improvements in transportation technology. The skill premium fell sharply in the 1910s. But, despite the even faster increase in the supply of high school skills during the 1920s, the skill premium levelled off and started a mild increase. Goldin and Katz (1995) conclude that the demand for high school graduates must have expanded sharply starting in the 1920s, presumably due to changes in office technology and higher demand from new industries such as electrical machinery, transport and chemicals (see also Goldin and Katz, 1998).[33]

Another interesting case study comes from the response of the Israeli labor market to the influx of large numbers of highly educated immigrants from the former Soviet Union. The size of this influx was enormous: migration increased the Israeli population by 12 percent in the first half of the 1990s. A theory with exogenous technology would predict a large decline in the relative wages of educated workers, very much as in the case of Palestinian labor discussed above.[34] In practice, the education premium did not fall (e.g., Friedberg, 1997). This seems to be mainly because most industries increased their employment of more skilled workers during this large influx (Gandal, Hanson, and Slaughter, 2000, and Cohen and Hsieh, 2000). This response suggests a change in the production structure towards more skilled workers, consistent with the theories outlined in this section.[35]

Despite this evidence showing simultaneous increases in the supply of, and demand for, skills in a number of episodes, it is difficult to distinguish exogenous and endogenous technical change. The exogenous technical change theory maintains that technical change is often skill-biased. Endogenous technical change theory instead suggests that new technologies should be skill-biased when the supply of skills increases. Since the supply of skills has increased most of the time over the past one hundred or so years, the implications of the two theories are quite similar. The increase in the supply of unskilled labor in the English cities during the early nineteenth century provides an interesting contrasting event for the two approaches. Bairoch (1988, p. 245) describes this rapid expansion as follows: "...between 1740 and 1840 the population of England...went up from 6 million to 15.7 million. ...while the agricultural labor force represented 60-70% of the total work force in 1740, by 1840 it represented only 22%." Habakkuk (pp. 136-137) also emphasizes the increase in the supply of unskilled labor in English cities, and attributes it to five sources. First, enclosures released substantial labor from agriculture. Second, "population was increasing

[33] As Goldin and Katz (2000) show using data from Iowa Prairies, returns to education were most likely higher in 1915 than in 1950. Although this evidence suggests that the long-run relative demand curve for skills was downward sloping over this period, it is consistent with the notion of skill-biased technical change induced by increased supply skilled workers. Specifically, during this period, demand for skills expanded very rapidly to accommodate the very large increase in the supply of high school graduates (see Goldin and Katz, 1995, 2000).

[34] The key difference between the two episodes is that Palestinian labor was a relatively small fraction of the Israeli work force, so we expect much less of a technology response to changes in the educational composition of Palestinian labor. Furthermore, Palestinian college graduates are not a close substitute for Israeli college graduates, and only a limited range of occupations are open to them.

[35] Since Israel can be approximated by a small open economy, another possibility is a change in the output mix and trade patterns. Gandal, Hanson, and Slaughter (2000) and Cohen and Hsieh (2000) find no evidence for this, and show that demand for skills increased in all Israeli sectors. Cohen and Hsieh (2000) also argue that because many Russian immigrants initially worked in low-skill occupations, the supply of skills to the Israeli economy may not have increased by much.

very rapidly" (p. 136). Third, labor reserves of rural industry came to the cities. Fourth, "there was a large influx of labor from Ireland" (p. 137). Finally, "technical changes in agriculture increased the supply of labor available to industry" (p. 137). Habakkuk further argues that this increase in the supply of labor was an important inducement to the development of factory methods. He quotes an American historian, Handlin, to explain why the adoption of factory methods in the U.S. were somewhat delayed. Handlin wrote:

> "no matter what degree of standardization technical process of manufacturing reached, the absence of a cheap labor supply precluded conversion to factory methods" (p. 146)

Habakkuk naturally placed much more importance on wage levels in determining innovation decisions, a view that later became known as the "Habakkuk thesis." But he also emphasized the different availabilities of skilled labor in Britain and the U.S.. He wrote:

> "in both countries this provided manufacturers with an incentive to adopt and devise methods which replaced skilled by non-skilled...[but]...the English had a stronger incentive than the Americans to replace skilled by unskilled labor." (p. 152)

With a similar reasoning to Habakkuk and Schmookler, the endogenous technology view suggests that businessmen took advantage of the rapid increase in the supply of labor in the cities by developing the factory system. According to this view, there is an intimate link between the skill-replacing technologies of the nineteenth century and the change in the factor supplies faced by employers—"the reserve army of unskilled labor". Although these historical examples are informative, they do not reveal whether endogenous technology choices are important in understanding more recent skill-biased technical change. A systematic study of how technologies respond to large changes in the relative supply of skills is clearly a worthwhile future research project.[36]

An important aspect of the endogenous technology theory is that it makes relatively tight predictions regarding the future path of technical change. While with exogenous skill bias theories, there is no clear reason to expect a further acceleration or deceleration in the skill bias of new technologies, according to this endogenous skill bias theory, the future path of technical progress should be closely tied to the future path of the supply of skills.

[36] It is also useful to note that the skill bias of technology probably responds not only to changes in the relative supply of skills, but to a variety of other factors. Recent work by Mobius (2000) and Thesmar and Thoning (2000) shows how the size of the product market, the degree of competitive pressure and instability facing firms may affect the way firms choose to organize, and therefore the demand for skills. Moreover, Mobius suggests that these changes reduced the demand for skills during the nineteenth century as there was greater standardization of products, but increased this demand during the past several decades as the need for flexibility increased. How the organization of product markets and the extent of competition affect technology choices and the demand for skills is a very promising area for future research.

If the relative supply of skills continues to increase, we should expect further skill-biased technical change.[37]

A major problem for the "technological revolution" models was the slowdown in TFP. Since the endogenous technology approach places the emphasis on which type of technologies are developed, it is not inconsistent with the slow growth in TFP during the past 30 years: an acceleration in the skill bias of new technologies does not require faster technical progress. The evidence presented in Newell, Jaffee and Stavins (1999) is consistent with the notion that changes in the direction of technical change can happen without an increase in the overall rate of productivity growth. They show that innovation in air conditioners responded to changes in energy prices by becoming more energy-efficient, but find no increase in the rate of productivity growth. In fact, Acemoglu (1998) and Kiley (1999) show that the increased effort of firms to develop more skill-biased technologies could run into decreasing returns, and hence may cause a slowdown in TFP growth. Intuitively, the overall productivity growth in the economy is maximized with a balanced distribution of resources towards developing skill-complementary and labor-complementary technologies (due to decreasing returns to each activity). During periods of rapid skill-biased technical change, all resources go into developing skill-complementary machines, and cause a sharp decline in advances in labor-complementary technologies. Because of the decreasing returns to scale, improvements in the other sector will not fully offset this decline, and overall TFP growth will fall.[38]

Finally, it is useful to discuss briefly the response of skills to technology. The analysis so far treated the supply of skills as exogenous, and investigated the implications of the supply on the demand for skills. Naturally, the supply of skills will also respond to economic incentives. In particular, more workers are likely to acquire skills when skill premia are higher. Such supply choices can be easily incorporated into this framework. Suppose, for example, that the relative supply of skills, H/L, is an increasing function of the skill premium, ω. In this case, if the long-run demand curve for skills is upward sloping, we can have an equilibrium path in which both the relative supply of skills and the skill premium increase together over time (see Acemoglu 1998). This equilibrium configuration gives us

[37] Although evidence from the 1990s suggests that skill-biased technical change is now slower, there is not yet sufficient evidence to decide whether the rapid skill-biased technical change of the 1980s has come to an end. It also has to be noted that the increase in the relative supply of college educated workers the 1990s may have been less than expected (see Figure 1), and this may have affected the technology choices of firms. In particular, with the increased labor force participation of less skilled workers, there may now be a sufficient number of unskilled workers supplying labor at low wages to make the further development of unskilled-labor-complementary technologies quite profitable.

[38] Interestingly, the view that too much effort towards improving the most skill-biased technologies may be related to the TFP slowdown is consistent with the pattern of sectoral TFP growth observed recently. As Gordon (1998) documents, there has been rapid TFP growth in computer producing sectors, but mediocre, or even disappointing, TFP growth in other sectors.

an attractive interpretation for the joint behavior of college skills and the college premium in the U.S. economy over the past sixty years (see Figure 1): large returns to schooling encourage education, which in turn induce more skill-biased technical change, increasing returns to schooling again.

5.3 A Puzzle: The Decline in the Wages of Low-Skill Workers

A common shortcoming of all the pure technology approaches discussed in this section is that they do not naturally predict stagnant average wages and/or falling wages for unskilled workers.[39] In the basic framework of Section 3, average wages should always increase when the supply of educated workers increases, and all wages should rise in response to an increase in the productivity of skilled workers, A_h. Yet, over the past 30 years the wages of low-skill workers have fallen in real value during, which contrasts with their steady increase in the 30 years previous.

Models of faster technological progress would naturally predict that unskilled workers should benefit from this faster progress. The endogenous technology approach discussed in the previous subsection, on the other hand, predicts that there may be no improvements in the technologies for unskilled workers for an extended period of time because skill-biased innovations are more profitable than labor-complementary innovations. Yet in that case, their wages should be stagnant, but not fall.

Some of the studies mentioned above have suggested explanations for the fall in the wages of low-skill workers. For example, recall that Galor and Moav (2000) argue that faster technological change creates an "erosion effect", reducing the productivity of unskilled workers. Using equation (3) from above, in the simplified version of their model discussed in Section 5.1, the wages of unskilled workers is $w_L = \phi_l(g)\, a\left[1 + \phi_h^{\rho}(H/L)^{\rho}\right]^{(1-\rho)/\rho}$, so the rate of growth of unskilled wages will be $\dot{w}_L/w_L = g\,(1 + \varepsilon_{\phi})$, where ε_{ϕ} is the elasticity of the ϕ_l function which is negative by the assumption that $\phi_l' < 0$. If this elasticity is less than -1, an acceleration in economic growth can reduce the wages of low-skill workers due to a powerful erosion effect.

Acemoglu (1999a) and Caselli (1999) derive a fall in the wages of less skilled workers because the capital-labor ratio for low education/low-skill workers falls as firms respond to technological developments. In Caselli's model this happens because the equilibrium rate of return to capital increases, and in my paper, this happens because firms devote more of their resources to opening specialized jobs for skilled workers.

Consider the following simple example to illustrate this point. There is a scarce supply of an input K, which could be capital, entrepreneurial talent or another factor of production.

[39] However, recall that if the increase in nonwage benefits are taken into account, average wages increased over this period. So the more robust fact might be the fall in the real wages of low skill workers.

Skilled workers work with the production function

$$Y_h = A_h^{\alpha} K_h^{1-\alpha} H^{\alpha} \tag{20}$$

while unskilled workers work with the production function

$$Y_l = A_l^{\alpha} K_l^{1-\alpha} L^{\alpha}, \tag{21}$$

where K_l and K_h sum to the total supply of K, which is assumed fixed. For simplicity, Y_l and Y_h are assumed to be perfect substitutes. In equilibrium, the marginal products of capital in two sectors have to be equalized, hence

$$\frac{K_l}{A_l L} = \frac{K - K_l}{A_h H}$$

Therefore, an increase in A_h relative to A_l will reduce K_l, as this scarce factor gets reallocated from unskilled to skilled workers. The wages of unskilled workers, $w_L = (1-\alpha) A_l^{\alpha} K_l^{1-\alpha} L^{\alpha-1}$, will fall as a result.

An innovative version of this story is developed by Beaudry and Green (2000). Suppose that equation (21) above is replaced by $Y_l = A_l^{\eta} K_l^{1-\eta} L^{\eta}$, with $\eta < \alpha$, and K is interpreted as physical capital. This implies that unskilled workers require more capital than skilled workers. Beaudry and Green show that an increase in H/L can raise inequality, and depress the wages of low-skill workers. Although this is related to the effects of the increase in the relative supply of skills on the path of technological progress discussed in the last subsection, the mechanism in Beaudry and Green's paper is quite different. The increase in H/L increases the demand for capital, and pushes the interest rate up. This increase in the interest rate hurts unskilled workers more than skilled workers because, given $\eta < \alpha$, unskilled workers are more "dependent" on capital.

A potential problem with both the Beaudry and Green and Caselli stories is that they explicitly rely on an increase in the price of capital. Although the interest rates were higher during the 1980s in the U.S. economy, this seems mostly due to contractionary monetary policy, and related only tangentially to inequality. Perhaps, future research will show a major role for the increase in the interest rates in causing the fall in the wages of low education workers over the past twenty-five years, but as yet, there is no strong evidence in favor of this effect.[40]

Overall, a potential problem for models based on technical change is to account for the decline in the wages of low-skill workers.[41] I argue in the next section that the effect

[40] Acemoglu (1999a), which is more in the spirit of the organizational theories discussed below, obtains the decline in the wages of unskilled workers through a change in the organization of production, which also entails a reallocation of capital away from them, but no increase in the rate of return to capital.

[41] Another possibility is that some of the technological developments of the past two decades have been "truly labor-replacing", for example, corresponding to an increase in $B_l(t)$ or b_t in terms of the production function in footnote 13. Autor, Levy and Murnane (2000), for example, suggest that computers have replaced unskilled routine tasks. This possibility has not been extensively researched yet.

of technical change on the organization of the labor market both amplifies the effect of technology on wage inequality, and provides an explanation for this decline.

6 Ramifications of Technical Change

This section discusses how technical change can affect labor market prices by transforming the organization of the labor market. The idea that technology affects the organization of production, and the institutions around it, is an old one. Marx put it in a dramatic fashion: "The hand-mill gives you society with the feudal lords; the steam-mill, society with the industrial capitalist." The argument here is not as extreme, but related: recent technological developments may have led to important changes in the organization of production.

My focus here is on three sets of changes that could account for the fall in the wages of low-skill workers: the transformation of the organization of firms; change in labor market institutions, particularly the decline in unionization; and the interaction between international trade and technical change. Organizational change often destroys the types of jobs that pay high wages to low-skill workers. Deunionization reduces the bargaining power of low-skill workers. And international trade with less developed countries (LDCs) increases the effective supply of unskilled labor and depresses the marginal value product of less skilled workers in the U.S. economy. Therefore, all three changes could be responsible for the decline in the wages of low-skill workers, and for the changes in the U.S. wage structure in general. Nevertheless, I argue that these factors by themselves are not the major cause of the increase in inequality. Instead they have become powerful actors only by interacting with technical change and have amplified the direct effect of technical change on inequality. They may even have caused the decline in unskilled wages.

6.1 Organizational Change and Inequality

A variety of evidence suggests that important changes in the structure of firms have been taking place in the U.S. economy over twenty-five years. Moreover, it seems clear that a major driving force for this transformation is changes in technologies (hence the view that technical change is essential for the changes in the organization of firms).

For example, team production and other high-performance production methods are now widespread in the U.S. economy (e.g., Ichinowski, Prennushi, and Show. 1997, or Applebaum and Batt, 1994). Similarly, Cappelli and Wilk (1997) show that there has been an increase in the screening of production workers, especially from establishments that use computer technology and pay high wages. Murnane and Levy (1996) report case study evidence consistent with this view. From their interviews with human resource personnel at a number of companies, they describe the change in the hiring practices of U.S. companies.

A manager at Ford Motor company in 1967 describes their hiring strategy as follows: "If we had a vacancy, we would look outside in the plant waiting room to see if there were any warm bodies standing there. If someone was there and they looked physically OK and weren't an obvious alcoholic, they were hired" (p. 19). In contrast, comparable companies in the late 1980s use a very different recruitment strategy. Murnane and Levy discuss the cases of Honda of America, Diamond Star Motors and Northwestern Mutual Life. All three companies spend substantial resources on recruitment and hire only a fraction of those who apply. Kremer and Maskin (1999) provide evidence of more segregation of workers across establishments. It seems that high wage workers are now much more concentrated in certain establishments. Similarly, in Acemoglu (1999a) I documented a change in the composition of jobs over the past 20 years. Figure 9 here replicates a pattern found in that paper, and plots the total percentage of workers employed in the top 25 percent and bottom 25 percent industry-occupation cells (what I called Weight-at-the-tails of the job quality distribution). These are the cells (job types) that pay relatively high or relatively low wages. In 1983, 35 percent of employment was in the top and bottom 25 percent job categories. By 1993, this number had risen to just under 38 percent. So, approximately 2.5 percent more workers now have either higher or lower quality jobs rather than medium quality jobs. The actual changes in the distribution of jobs may be much larger than this, since substantial changes in the types of jobs often take place within given occupations.

The view that changes in the organization of firms have had a fundamental effect on the labor market is often expressed in the popular press, and in the organizational literature (e.g., Zuboff, 1988). The first paper to formalize this organizational approach is Kremer and Maskin (1999), followed by Acemoglu (1999a), Mobius (2000) and Thesmar and Thoning (2000). Kremer and Maskin consider a production function which distinguishes between managers and workers. They show that a change in technology or an increase in the dispersion of skills may encourage high skill workers to match with other high skill workers, rather than work as managers in establishments employing low-skill workers.

Here I outline a simple model, inspired by Kremer and Maskin (1999) and Acemoglu (1999a), that captures the effect of the changes in technologies on the organization of production. The basic idea is simple. As the productivity of skilled workers increases, it becomes more profitable for them to work by themselves in separate organizations rather than in the same workplace as unskilled workers. This is because when the skilled and unskilled work together, their productivities interact, and unskilled workers may put downward pressure on the productivity of skilled workers.

Specifically, suppose that firms have access to the following production functions

$$\begin{aligned} \text{the old-style production function} &: \quad Y = B_p \left[(A_l L)^\rho + (A_h h_O)^\rho \right]^{1/\rho}, \\ \text{the new-organization production function} &: \quad Y = B_s A_h^\beta h_N. \end{aligned}$$

Intuitively, skilled and unskilled workers can either be employed in the same firm as with the old-style function, h_O, or high skill workers can be employed in separate firms, h_N. The fact that when they are employed in the same firm, these two types of workers affect each other's productivity is captured by the CES production function. This formulation implies that if the productivity (ability) of unskilled workers, A_l, is very low relative to A_h, they pull down the productivity of skilled workers. In contrast, when they work in separate firms, skilled workers are *unaffected* by the productivity of unskilled workers. Moreover, $\beta > 1$, which implies that improvements in the productivity of skilled workers has more effect on the productivity of new style organizations. The parameters B_p and B_s capture the relative efficiency of old and new style production functions.

The labor market is competitive, so the equilibrium organization of production will maximize total output, given by $B_p\left[(A_l L)^\rho + (A_h h_O)^\rho\right]^{1/\rho} + B_s A_h^\beta (H - h_O)$, where $h_O \in [0, H]$ is the number of skilled workers employed in the old-style organizations. For all cases in which $h_O > 0$, the solution to this problem will involve

$$w_H = B_p A_h^\rho h_O^{\rho-1}\left[A_l^\rho L^\rho + A_h^\rho h_O^\rho\right]^{(1-\rho)/\rho} = B_s A_h^\beta, \tag{22}$$

i.e., skilled workers need to be paid $B_s A_h^\beta$ to be convinced to work in the same firms as the unskilled workers. The unskilled wage is

$$w_L = B_p A_l^\rho L^{\rho-1}\left[A_l^\rho L^\rho + A_h^\rho h_O^\rho\right]^{(1-\rho)/\rho} < w_H \tag{23}$$

Now consider an increase in A_h. Differentiating (22) yields $\partial(A_h h_O)/\partial A_h < 0$, which, from (23), implies that $\partial w_L/\partial A_h < 0$. Therefore, skill-biased technical change encourages skilled workers to work by themselves, and as a result, unskilled wages fall. Intuitively, since, in the old-style organizations, the productivity of skilled workers depends on the ability of unskilled workers, when the skilled become even more productive, the downward pull exerted on their productivity by the unskilled workers becomes more costly, and they prefer to work in separate organizations. This reduces the ratio h_O/L and depresses unskilled wages. As a result, improvements in technology, which normally benefit unskilled workers as in Section 3, may actually hurt unskilled workers because they transform the organization of production.

An increase in B_s/B_p, which raises the relative profitability of the new organizational form, also leads to further segregation of skilled and unskilled workers in different organizations. This last comparative static result is useful since Bresnahan (1999) and Autor, Levy and Murnane. (2000) argue that by replacing workers in the performance of routine tasks, computers have enabled a radical change in the organization of production.[42] This is

[42] A related perspective is offered by Aghion (2000), who also argues that computers replace unskilled tasks. He suggests that computers are a "general-purpose technology", so their diffusion follows an inverted S shaped pattern, and as more and more firms adopted computers, demand for unskilled workers fell rapidly.

reminiscent to a technological change that makes the new-organization production function more profitable.

These organizational stories are attractive since they provide a unified explanation for the changes in the wage structure and the apparent changes in the organization of firms. An interesting recent paper by Caroli and Van Rennan (1999) provides evidence suggesting that changes in wages have been accompanied by changes in organizational forms. Acemoglu (1999a) and Kremer and Maskin (1999) also provide evidence suggesting a number of organizational changes in the U.S. economy during the past 25 years. Nevertheless, this evidence does not yet enable an assessment of whether changes in organizational forms have been an important contributor to the changes in labor market prices, and future research is required to determine the role of organizational change.

6.2 Institutional Change

Two major changes in labor market institutions over the past twenty five years are the decline in the real value of state and federal minimum wages and the reduced importance of trade unions in wage determination. Many economists suspect that these institutional changes may be responsible for the changes in the structure of the U.S. labor market (see Freeman, 1991, DiNardo, Fortin and Lemieux, 1995, Lee, 1999).

The real value of the minimum wage eroded throughout the 1980s as nominal minimum wages remained constant for much of this period. Since minimum wages are likely to increase the wages of low paid workers, this decline may be responsible for increased wage dispersion. DiNardo et al. (1995) and Lee (1999) provide evidence in support of this hypothesis. Although the contribution of minimum wages to increased wage dispersion cannot be denied, minimum wages are unlikely to be a major factor in the increase of overall inequality. First, only a very small fraction of male workers are directly affected by the minimum wage (even in 1992, after the minimum wage hike of 1990-91, only 8 percent of all workers between the ages of 18 and 65 were paid at or below the minimum wage). Although minimum wages may increase the earnings of some workers who are not directly affected, they are highly unlikely to affect the wages above the median of the wage distribution. But as Figure 3 shows, the difference between the 90th percentile and the median mirrors the behavior of the difference between the median and the 10th percentile.[43] This implies that whatever factors were causing increased wage dispersion at the top of the distribution are likely to have been the major cause of the increase in wage dispersion throughout the distribution. Second, the erosion in the real value of the minimum wage started in the 1980s, whereas, as shown above, the explosion in overall wage inequality

[43]Except during the early 1980s when there is a more rapid increase in inequality at the bottom of the wage distribution, most likely due to the falling real value of the minimum wage.

began in the early 1970s.

The declining importance of unions may be another important factor in the increase in wage inequality. Unions often compress the structure of wages and reduce skill premia (see, for example, Reynolds, 1951, or Freeman and Medoff, 1984). Throughout the post-war period in the U.S. economy, unions negotiated the wages for many occupations, even indirectly influenced managerial salaries (see DiNardo, Hallock and Pischke, 2000). Unions also explicitly tried to compress wage differentials. This suggests that the decline of unions may have been a leading cause of the changes in the structure of wages.

Although deunionization could in principle be an important factor in the structure of wages, the extent and timing of deunionization suggests that it is not the major driving force of the increase in inequality. First, wage inequality increased in many occupations in which prices were never affected by unions (such as lawyers and doctors). Moreover, in the U.S., deunionization started in the 1950s, a period of stable wage inequality. During the 1970s, though unionization fell in the private sector, overall unionization rates did not decline much because of increased unionization in the public sector. Overall union density was approximately constant, around 30 percent of the work force, between 1960 and 1975. It was the anti-union atmosphere of the 1980s and perhaps the defeat of the Air-traffic Controllers' Strike that led to the most major declines of the unions, dating the sharp declines in unionization after the rapid increase in inequality during the early 1970s.[44] Evidence from other countries also paints a similar picture. For example, in the UK, wage inequality started its sharp increase in the mid 1970s, while union density increased until 1980 and started the rapid decline only during the 1980s (Gosling, 1998). In Canada, while unionization rates increased from around 30 percent in the 1960s to over 36 percent in the late 1980s (Riddell, 1993, table 4.1), wage inequality also increased (see, for example, Freeman and Needels, 1993, figure 2.4).

Although the timing of deunionization is inconsistent with this institutional change being the major cause of the change in wage inequality, deunionization could clearly affect the wages of unskilled workers. One possibility is that deunionization is a contributing factor. But why did the unions decline while technology was rapidly becoming more skill-biased? Acemoglu, Aghion and Violante (2000) suggest that deunionization may have been caused by the technological developments of the past decades. According to this perspective, technical change is the driving force of the changes in the wage structure, but it also causes deunionization, and the resulting deunionization can have a very large indirect effect on wage inequality, and cause the wages of less skilled workers to fall.

To see the basic argument suppose that production can be carried out either in unionized or nonunionized firms. In nonunionized firms, workers receive their marginal products,

[44] An interesting recent paper Farber and Western (2000) date the major decline in union activity to the early 1980s, a few months before the Air-traffic Controllers Strike.

which I denote by A_h and A_l for skilled and unskilled workers. Assume that unions always compress the structure of wages—i.e., they reduce wage differentials between skilled and unskilled workers. This wage compression could be driven by a variety of factors. Acemoglu, Aghion and Violante (2000) argue that unions encourage productive training, and such training is incentive compatible for firms only when the wage structure is compressed (see also Acemoglu and Pischke, 1999). Alternatively, collective decision making within a union may reflect the preferences of its median voter, and if this median voter is an unskilled worker, he will try to increase unskilled wages at the expense of skilled wages. It is also possible that union members choose to compress wages because of ideological reasons or for social cohesion purposes. The empirical literature supports the notion that unions compress wages, though it does not distinguish among the various reasons (see Freeman and Medoff, 1984). I capture this in a reduced form way here using the equation

$$\omega = \frac{w_H}{w_L} \leq \psi \frac{A_h}{A_l}, \tag{24}$$

where $\psi < 1$. Unions could never attract skilled workers unless they provide some benefits to them. Here I simply assume that they provide a benefit β to all workers, for example, because unions increase productivity (e.g., Freeman and Medoff, 1984, and Freeman and Lazear, 1995), or because they encourage training. Alternatively, β could be part of the rents captured by the union. The zero profit constraint for firms would be: $(w_H - \beta)H + (w_L - \beta)L \leq A_h H + A_l L$ (in the case where β stands for rents, the zero profit ensures that the firm does not wish to open a non-union plant). Combining this equation with (24), and assuming that both hold as equality, we obtain

$$w_H = \frac{(A_h + \beta)\, H + (A_l + \beta)\, L}{A_h H + \psi^{-1} A_l L} A_h. \tag{25}$$

Skilled workers will be happy to be part of a union as long as w_H given by (25) is greater than A_h. As A_h/A_l increases—i.e., as there is further skill-biased technical change—, w_H will fall relative to A_h. Therefore, skill-biased technical change makes wage compression more costly for skilled workers, eventually destroying the coalition between skilled and unskilled workers that maintains unions.

The important point is that deunionization causes a decline in the wages of unskilled workers from $w_L = \frac{(A_h+\beta)H+(A_l+\beta)L}{A_h H+\psi^{-1}A_l L}\psi A_l$ to A_l. Unskilled workers, who were previously benefiting from wage compression imposed by unions, experience a fall in real earnings as a result of deunionization. Therefore, in this model, technical change not only affects wage inequality directly, but also induces a change in labor market institutions. The effect of this change in institutions on inequality can be potentially larger than the direct effect of technical change, and explain the decline in the real wages of less skilled workers.

Although Acemoglu, Aghion and Violante (2000) provide some evidence consistent with these patterns, whether deunionization was important in the decline of the wages of low-skill workers, and whether technical change is responsible for deunionization are open questions. It might also be interesting to investigate whether changes in technology may have also affected the coalition supporting minimum wages, and hence played a role in the decline in the minimum wage during the 1980s. For example if unions, which have traditionally supported minimum wages, were weakened by the technological developments of the decade as argued in this subsection, technological developments may have indirectly contributed to the weakening of the coalition in support of minimum wages.[45]

6.3 Trade, Technical Change and Inequality

Finally, I discuss another major change affecting the U.S. economy: the increased volume of international trade. Increased international trade by itself is not the cause of the changes in the U.S. wage structure, but trade could be a very important factor in determining wage inequality if it also affects the direction of technical change. Therefore, this subsection is somewhat different from the previous two, since it is not about the effect of technology on the organization of the labor market, but on the effect of a major change in regulations on technological development. I group it with the other changes since it is about interactions between technical change and the organization of the labor market, and since it provides a possible explanation for the fall in the wages of the less skilled workers.

Standard trade theory predicts that increased international trade with less developed countries (LDCs), which are more abundant in unskilled workers, should increase the demand for skills in the U.S. labor market. Therefore, the increase in international trade may have been the underlying cause of the changes in U.S. wage inequality.

To discuss these issues, consider the two good interpretation of the model in Section 3. Consumer utility is defined over $[Y_l^\rho + Y_h^\rho]^{1/\rho}$, with the production functions for two goods being $Y_h = A_h H$ and $Y_l = A_l L$. Both goods are assumed to be tradable. For simplicity, let me just compare the U.S. labor market equilibrium without any trade to the equilibrium with full international trade without any trading costs.

Before trade, the U.S. relative price of skill intensive goods, p_h/p_l, is given by

$$p^{US} = \frac{p_h}{p_l} = \left[\frac{A_h H}{A_l L}\right]^{\rho-1}. \tag{26}$$

[45] Finally, both the decline in the role that unions play and in the value of minimum wages may have been caused by changes in social norms, which could also be responsible for the increase in inequality (e.g., the emergence of the winner-take-all society as claimed by Frank and Cook, 1996). Unfrotunately, there is currently little research on the effect of social norms on inequality and on why inequality norms may have changed over the past 30 years.

The skill premium is then simply equal to the ratio of the marginal value products of the two types of workers, that is, $\omega^{US} = p^{US} A_h / A_l$. Next, suppose that the U.S. starts trading with a set of LDCs that have access to the same technology as given by A_h and A_l, but are relatively scarce in skills. Denote the total supplies of skilled and unskilled workers in the LDCs by $\widehat{H}$ and $\widehat{L}$ where $\widehat{H}/\widehat{L} < H/L$, which simply reiterates that the U.S. is more abundant in skilled workers than the LDCs.

After full trade opening, the product markets in the U.S. and the LDCs are joined, so there will be a unique world relative price. Since the supply of skill and labor-intensive goods is $A_h \left(H + \widehat{H}\right)$ and $A_l \left(L + \widehat{L}\right)$, the relative price of the skill intensive good will be

$$p^W = \left[\frac{A_h \left(H + \widehat{H}\right)}{A_l \left(L + \widehat{L}\right)} \right]^{\rho - 1} > p^{US}. \tag{27}$$

The fact that $p^W > p^{US}$ follows immediately from $\widehat{H}/\widehat{L} < H/L$. Intuitively, once the U.S. starts trading with skill-scarce LDCs, demand for skilled goods increases and pushes the prices of these goods up.

Labor demand in this economy is derived from product demands. The skill premium therefore follows the relative price of skill-intensive goods. After trade opening, the U.S. skill premium increases to

$$\omega^W = p^W \frac{A_h}{A_l} > \omega^{US} \tag{28}$$

where the fact that $\omega^W > \omega^{US}$ is an immediate consequence of $p^W > p^{US}$. Therefore, trade with less developed countries increases wage inequality in the U.S..

The skill premium in the LDCs will also be equal to ω^W after trade since the producers face the same relative price of skill-intensive goods, and have access to the same technologies. Before trade, however, the skill premium in the LDCs was $\widehat{\omega} = \widehat{p} A_h / A_l$, where $\widehat{p} = \left(A_h \widehat{H} / A_l \widehat{L} \right)^{\rho - 1}$ is the relative price of skill-intensive goods in the LDCs before trade. The same argument as above implies that $\widehat{p} > p^W$, i.e., trade with the skill-abundant U.S. reduces the relative price of skill-intensive goods in the LDCs. This implies that $\omega^W < \widehat{\omega}$; after trade wage inequality should fall in the LDCs.

Although in theory increased trade with the LDCs can be the cause of the rapid increase in the demand for skills, most evidence suggests that the direct effect of increased international trade on the U.S. labor market has been relatively minor.

First, as equation (27) shows, the effect of international trade works through a *unique intervening mechanism*: more trade with the LDCs increases the relative price of skill-intensive goods, p, and affects the skill premium via this channel. In fact, in this simple framework, the percentage increase in the skill premium is directly proportional to the percentage increase in the relative price of skill-intensive goods. Perhaps the most damaging piece of evidence for the trade hypothesis is that most studies suggest the relative price of skill-intensive goods did not increase over the period of increasing inequality. Lawrence and Slaughter (1993) found that during the 1980s the relative price of skill-intensive goods actually fell. Sachs and Shatz (1994) found no major change or a slight decline. A more recent paper by Krueger (1997) criticized the methods and data used by these studies, and found an increase in the relative price of skill intensive goods. Nevertheless, the increase in these prices is relatively small, so would not be able to account for the large increase in the skill premium experienced in the U.S. economy (recall that the change in the relative price of skill-intensive goods needs to be of the same order of magnitude as the change in the skill premium).

Second, with trade as the driving force, increased production of skill-intensive goods should be drawing workers away from other sectors. In contrast, as documented by Murphy and Welch (1993), Berman, Bound and Griliches (1994) and Autor, Katz and Krueger (1998), all sectors, even those producing less skilled goods, increased their demands for more educated workers. This pattern is not consistent with trade being the main driving force of the increase in the demand for skilled workers (though one has to bear in mind increased outsourcing in interpreting this fact, see Feenstra and Hanson, 1999).

Third, a direct implication of the trade view is that, as shown above, while demand for skills and inequality increase in the U.S., the converse should happen in the LDCs that have started trading with the more skill abundant U.S. economy. The evidence, however, suggests that more of the LDCs experienced rising inequality after opening to international trade (see Hanson and Harrison, 1994, or Robbins, 1995). Although the increase in inequality in a number of cases may have been due to concurrent political and economic reforms, the preponderance of evidence is not favorable to this basic implication of the trade hypothesis.

Finally, a number of economists have pointed out that U.S. trade with the LDCs is not important enough to have a major impact on the U.S. product market prices and consequently on wages. Krugman (1995) illustrates this point by undertaking a calibration of a simple North-South model. Katz and Murphy (1992), Berman, Bound and Griliches (1994) and Borjas, Freem and and Katz (1997) emphasize the same point by showing that the content of unskilled labor embedded in the U.S. imports is very small relative to the changes in the supply of skills taking place during this period.[46]

[46] This is probably the weakest criticism against the trade view, and many studies have pointed out how international trade could have a larger effect on U.S. labor market prices in the presence of labor market rents. For example, Borjas and Ramey (1995), Rodrik (1996) and Dube and Reddy (1999) have argued that the threat of international trade may reduce wages, especially in sectors with substantial rents, and this change in bargaining power may affect the earnings of unskilled workers more, increasing inequality.

Although the above arguments suggest that increased international trade with LDCs is not the major cause of the changes in the wage structure by itself, they do not rule out a powerful effect of international trade when it interacts with technical change. In particular, in a world with endogenous technical change, increased international trade could affect the types of technologies developed and adopted by firms, and have a large effect through this channel. This possibility was first raised by Adrian Wood (1994) who argued that trade with the LDCs will lead to *defensive skill-biased innovations*. Wood, however, did not develop the mechanism through which such defensive innovations could occur. I now illustrate how trade causes skill-biased technical change using the endogenous technology model developed in Section 5.2 (this analysis draws on Acemoglu, 1999b).

Suppose that the U.S. starts trading with the LDCs as discussed above, and assume that the LDCs always use U.S. technologies. Therefore, the supply of skilled and unskilled goods in the LDCs is $Y_h = A_h \widehat{H}$ and $Y_l = A_l \widehat{L}$ where as before $\widehat{H}/\widehat{L} < H/L$. The immediate effect will be an increase in the relative price of skill-intensive goods as illustrated by equation (27). Now, recall from the analysis in Section 5.2 that there is a relative price effect on the direct technical change: developing technologies to produce the more expensive good is more profitable. Trade, by making the skill-intensive goods more expensive, encourages more skill-biased technical change.

To determine exactly how the direction of technical change will be affected by trade, we need to know the market sizes for new technologies after trade opening. It is plausible to assume that trade opening with the LDCs will not have a major effect on the enforcement of intellectual property right in the South. In that case, trade opening will induce skill-biased technical change in the U.S.. Specifically, as long as after trade opening the U.S. does not start producing technologies for unskilled workers in the LDCs, the relative market sizes for the two types of technologies remain at H/L. This implies that the technology market clearing condition, equation (18), no longer holds. In particular, since $p_h^W H > p_l^W L$ from equation (27), there will now only be incentives for skill-biased technical change. This process will continue until equation holds again, i.e., until $p^W = p_h^W / p_l^W = (H/L)^{-1}$. Therefore, the direction of technical change is still determined by equation (18) from Section 5.2, i.e., by U.S. domestic relative supplies alone. Intuitively, the relative price of skill-intensive goods plays two roles in this model. The first is to clear the market for goods (i.e., equation (27)), and the second is to ensure equilibrium in the technology market, as captured by equation (18). Since the technology market clearing condition relates the relative price of skill-intensive goods to the relative supplies in the U.S. market, which do not change, the long-run equilibrium price of skill-intensive goods cannot change either.

Combining equations (18) and (27) gives

$$\frac{A_h^W}{A_l^W} = \left[\frac{\left(H+\widehat{H}\right)}{\left(L+\widehat{L}\right)}\right]^{-1}\left(\frac{H}{L}\right)^{1/(1-\rho)} > \frac{A_h^{US}}{A_l^{US}} = \left(\frac{H}{L}\right)^{\rho/(1-\rho)}$$

where A_h^{US}/A_l^{US} is the pretrade skill bias of technology in the U.S..

The implication is that when the direction of technical change in endogenous, trade between the U.S. and the LDCs will induce skill-biased technical progress. The result is not only that trade leads to an increase in skill premia, but that this can happen without the counterfactual implications of the standard trade models discussed above

The criticisms levied against the standard trade model do not apply in this model with endogenous skill bias, mostly because trade with the LDCs induces skill-biased technical change. The first implication of this induced skill bias is that the impact of trade on labor markets may be much larger than predicted by the standard trade models, which helps against the criticism that the amount of trade the U.S. undertakes with the LDCs is not large enough. Second, because trade causes skill-biased technical change, the fact that all sectors have increased the employment of skilled workers is consistent with trade being the underlying cause of the increase in inequality. Third, for the same reason, there is a force counteracting the decline in inequality in the LDCs implied by trade: these economies use U.S. technologies, which are becoming more skill-biased. Finally, and quite strikingly, trade leaves the relative price of skill-intensive goods in the U.S. unchanged in the long-run. Recall that changes in relative prices are the usual intervening mechanism in trade models, so a number of studies have concluded that trade has not been an important factor in the increase in inequality because the relative price of skill-intensive goods has not increased much (e.g., Lawrence and Slaughter, 1994, Sachs and Shatz, 1995). In this model, however, we should expect no such changes; the long-run relative price of skill-intensive goods in the U.S. is unaffected by trade. More generally, induced skill-biased technical change in the U.S. implies that trade will increase the price of skill-intensive goods by only a limited amount, but will still have a major effect on the U.S. labor market.

Notice finally that the interaction between international trade and technical change may help to explain the decline in the wages of low-skill workers. Increased international trade acts as an increase in the supply of unskilled workers, and as shown in our basic framework in Section 3, puts downward pressure on unskilled wages. Borjas, Freeman and Katz (1987), for example, provide evidence that increased international trade during the 1980s reduced the wages of high school graduate workers (though they also suggest that the effect of the immigration of less skilled workers was greater). So overall, international trade could still be a major driving force of the changes in the wage structure. However, for increased trade to have such a large effect on the structure of wages—and to avoid the aforementioned counterfactual implications— it must cause a change in the path of technological progress. Therefore, what we have here is an alternative theory of acceleration: skill-biased technical change accelerates, neither exogenously nor in response to the changes in the supply of skilled workers, but in response to change in relative prices due to trade opening.

7 Changes in residual inequality

The previous sections highlighted that there are major unanswered questions regarding the causes of the increase in inequality, the reasons for the faster skill-biased technical change in the past few decades, and the determinants of the fall in the wages of low-skill workers. These big questions awaiting further research not withstanding, we have a reasonably simple and useful framework, and the beginnings of consistent answers. In contrast, in this and the next section, I discuss areas where answers are much more tentative, and there is much uncertainty and need for further research.

I begin with residual inequality. A major issue that most models discussed so far failed to address is the differential behavior of returns to schooling and residual inequality during the 1970s. I argue in this section that an explanation for this pattern requires models with multi-dimensional skills.

7.1 A single index model of residual inequality

The simplest model of residual inequality is a single index model, in which there is only one type of skill, though this skill is only imperfectly approximated by education (or experience). Expressed alternatively, in a single index model observed and unobserved skills are perfect substitutes. Consider, for example, the model developed above, but suppose that instead of skills, we observe education, e.g. whether the individual is a college graduate, which is imperfectly correlated with skills. A college graduate has a probability ϕ_c of being highly skilled, while a noncollege graduate is high skill with probability $\phi_n < \phi_c$. Suppose that the skill premium is $\omega = w_H/w_L$. The college premium in this case is

$$\omega^c = \frac{w_C}{w_N} = \frac{\phi_c w_H + (1-\phi_c)\, w_L}{\phi_n w_H + (1-\phi_n)\, w_L} = \frac{\phi_c \omega + (1-\phi_c)}{\phi_n \omega + (1-\phi_n)},$$

while within-group inequality, i.e., the difference between high wage college graduates (or noncollege graduates) and low-wage college graduates (or noncollege graduates), is $\omega^{within} = \omega$. It is immediately clear that both ω^c and ω^{within} will always move together—as long as ϕ_c and ϕ_n remain constant. Therefore, an increase in the returns to observed skills—such as education— will also be associated with an increase in the returns to unobserved skills.

This framework provides a natural starting point, linking between and within-group inequality, but it predicts that within and between-group inequality should move together.

However, as discussed above, during the 1970s, returns to schooling fell while residual group inequality increased sharply. We can only account for this fact by positing a decline in ϕ_c relative to ϕ_n of a large enough magnitude to offset the increase in ω; this would ensure that during the 1970s the college premium could fall despite the increase in within group inequality. A large decline in ϕ_c relative to ϕ_n would predict a very different behavior of the college premium within different cohorts, but the Appendix shows little evidence in favor of this. I therefore conclude that the single index model cannot explain the changes in residual inequality during the 1970s and 1980s.

7.2 Sorting and residual inequality

Another approach would combine educational sorting with an increase in the demand for skills. Suppose, for example, wages are given by $\ln w_{it} = \theta_t a_i + \gamma_t h_i + \varepsilon_{it}$ where h_i is a dummy for high education, a_i is unobserved ability, and ε_{it} is a mean zero disturbance term. Here γ_t is the price of observed skills, while θ_t is the price of unobserved skills. The education premium can be written as

$$\ln \omega_t \equiv E\left(\ln w_{it} \mid h_{it} = 1\right) - E\left(\ln w_{it} \mid h_{it} = 0\right) = \gamma_t + \theta_t (A_{1t} - A_{0t})$$

where $A_{1t} \equiv E\left(\ln w_{it} \mid h_{it} = 1\right)$ and A_{0t} is defined similarly. Residual (within-group) inequality can be measured as $Var\left(A_{it} \mid h = 0\right)$ and $Var\left(A_{it} \mid h = 1\right)$.

Under the assumption that there is perfect sorting into education, i.e., that there exists a threshold $\overline{a}$ such that all individuals with unobserved ability $\overline{a}$ obtain education, within-group inequality among high and low education workers will move in opposite directions as long as the price of observed skills, θ, is constant. To see this, note that when θ is constant and $\overline{a}$ declines (i.e., average education increases), $Var\left(A_{it} \mid h = 1\right)$ will increase, but $Var\left(A_{it} \mid h = 0\right)$ will fall. Intuitively, there are more and more "marginal" workers added to the high education group, creating more unobserved heterogeneity in that group and increasing within-group inequality. But in contrast, the low education group becomes more homogeneous. Therefore, without a change in the prices for unobserved skills, this approach cannot account for the simultaneous increase in inequality both among low and high education groups.

A natural variation on this theme would be a situation in which γ and θ move together. However, this will run into the same problems as the single index model: if γ and θ always move together, then such a model would predict that within-group inequality should have fallen during the 1970s. Therefore, models based on sorting also require a mechanism for the prices of observed and unobserved skills to move differently during the 1970s.

7.3 Churning and residual inequality

Another approach emphasizes that workers of all levels of education may face difficulty adapting to changes. This has been argued by Aghion, Howitt and Violante (1999) and Gould, Moav and Weinberg (2000). According to this approach, an increase in inequality also results from more rapid technical change, not because of skill bias but because of increased "churning" in the labor market. Aghion, Howitt and Violante (1999), for example, suggest that only some workers will be able to adapt to the introduction of new technology, and this will increase wage inequality.

An advantage of this approach is that it is in line with the increased earnings instability pointed out by Gottschalk and Moffit (1994). However, there is relatively little evidence other than this increase in earnings instability that supports the notion that there is more churning in the labor market. The data on job creation and job destruction reported by Davis, Haltiwanger and Schuh (1996) shows no increase in job reallocation during the 1980s or early 1990s, and most evidence does not indicate much of a decline in job stability over this period (e.g. Diebold et al., 1997, or Farber, 1995).[47] Also theories based on churning do not naturally predict a divergence between returns to educations and residual inequality during the 1970s. Therefore, a mechanism that could lead to differential behavior in the prices to observed and unobserved skills is still necessary.[48]

7.4 A two-index model of residual inequality

Since models based on a single index of skill (or models where different types of skills are perfect substitutes) are inconsistent with the differential behavior of returns to schooling and within-group inequality during the 1970s, an obvious next step is to consider a two-index model where observed and unobserved skills are imperfect substitutes (see Acemoglu, 1998). In particular, suppose that there are four types of workers, differentiated by both

[47] More recent evidence indicates that there may have been a decrease in job tenure during the later parts of the 1990s (see, Neumark, et al, 1999).

[48] An interesting theory similar to the churning models that could lead to such a differential behavior is advanced by Galor and Tsiddon (1997). They draw a distinction between ability and education, and argue that returns to ability increase faster during periods of rapid technological change (see also Galor and Moav, 2000). If we view the 1970s as a period of rapid technological change, as suggested above, this theory would imply an increase in the returns to ability (unobserved skills) during this period. Nevertheless, this interesting explanation is still not consistent with the facts since it predicts that returns to education should have also increased during the 1970s, though less than residual inequality. This is because high ability individuals are more likely to be high education, so rapid technological progress should also increase returns to schooling. Perhaps a combination of this mechanism with differential sorting into education, or with imperfect substitution between high and low education workers, might be able to account for the divergence between returns to schooling and residual inequality during the 1970s, but such a model has not been developed yet.

education and unobserved skills. The economy has an aggregate production function

$$Y = \left[(A_{lu}L_u)^\rho + (A_{ls}L_s)^\rho + (A_{hu}H_u)^\rho + (A_{hs}H_s)^\rho\right]^{1/\rho},$$

where L_u is the supply of low-skill low education workers, and other terms are defined similarly. Within-group inequality corresponds to the ratio of the wages of skilled low education workers to those of unskilled low education workers, and/or to the ratio of the wages of skilled high education workers to those of unskilled high education workers. A natural starting point is to presume that the fraction of high skill workers in each education group is constant, say at $\phi_l = L_s/L_u$ and $\phi_h = H_s/H_u > \phi_l$, which implies that there are more high ability workers among high education workers. With this assumption, within-group inequality measures will be

$$\frac{w_{Ls}}{w_{Lu}} = \left(\frac{A_{ls}}{A_{lu}}\right)^\rho \phi_l^{-(1-\rho)} \text{ and } \frac{w_{Hs}}{w_{Hu}} = \left(\frac{A_{hs}}{A_{hu}}\right)^\rho \phi_h^{-(1-\rho)}. \tag{29}$$

The college premium, on the other hand, is

$$\omega = \frac{\phi_h^\rho A_{hs}^\rho + A_{hu}^\rho}{\phi_l^\rho A_{ls}^\rho + A_{lu}^\rho}\left(\frac{1+\phi_l}{1+\phi_h}\right)^\rho \left(\frac{H}{L}\right)^{-(1-\rho)}.$$

Using this framework and the idea of endogenous technology, we can provide an explanation for the differential behavior of returns to schooling and within-group inequality during the 1970s. Recall that according to the endogenous technology approach, it is the increase in the supply of more educated workers that triggers more rapid skill-biased technical change. Because technology adjusts sluggishly, the first effect of an increase in the supply of educated workers, as in the 1970s, will be to depress returns to schooling, until technology has changed enough to offset the direct effect of supplies (see Figure 7). This change in returns to schooling has no obvious implication for within-group inequality in a multi-skill set up since it is the education skills that are becoming abundant, not unobserved skills—in fact in equation (29) within-group inequality is invariant to changes in the supply of educated workers unless there is a simultaneous change in ϕ_h and ϕ_l.

Under the plausible assumption that more skilled workers within each education group also benefit from skill-biased technical progress, technical change spurred by the increase in the supply of educated workers will immediately start to benefit workers with more unobserved skills, raising within-group inequality. Therefore, an increase in the supply of educated workers will depress returns to schooling, while increasing within-group inequality. After this initial phase, technical change will increase both returns to schooling and within-group inequality.[49]

[49]If the 1960s are also characterized by steady skill-biased technical change, equation (29) suggests that

Overall, although single index models are not capable of explaining the changes in residual inequality over the past thirty years, we do not know how important other factors are. Analysis of the determinants of residual inequality and the reasons why there was an explosion in overall inequality beginning in the 1970s remains a major research area.

8 Cross-country patterns

So far, I have focused on U.S. wage inequality patterns and incentives to develop new technologies coming from the U.S. supply of skills. The cross-country dimension presents a number of challenges. First, it is difficult to explain why inequality increased much more in some countries than others. Second, when there are many countries in the world economy, is it the relative supply of skills in each country or in the world as a whole that determines the direction of technical change? I now briefly discuss these issues.

8.1 Differences in inequality patterns

Although the tendency towards greater inequality has been a feature in many developed and LDCs (see Freeman and Katz 1995, and Berman and Machin, 2000), there are also marked differences in the behavior of within and between-group inequality across these countries. Katz, Blanchflower, and Loveman (1995) and Murphy, Riddell, and Romer (1998) show that the differential behavior of the supply of skills can go a long way towards explaining the differences in the returns to schooling, especially between the U.S., Canada and the U.K. Nevertheless, it is puzzling that wage inequality increased substantially in the U.S. and the UK, but remained fairly stable in many continental European economies (see, for example, Davis, 1995, Gottschalk and Smeeding, 1999).

The standard explanation for this divergent behavior, succinctly summarized by Krugman (1994) and OECD (1994), and sometimes referred to as the Krugman hypothesis, maintains that inequality did not increase as much (or not at all) in Europe because labor market institutions there encourage wage compression, limiting the extent of inequality. This can be captured in the competitive framework of Section 3, where firms are always along their relative demand curve, by assuming that labor market institutions impose an

there should have been a steady increase in residual inequality during this decade as well. The data presented in Section 2 do not support this prediction. Therefore, it seems that to explain the basic trends, one needs to posit that improvements in technology take the form $\ln A_j(t) = \gamma'_j + \gamma_j \cdot t$ with $\gamma_{hs} = \gamma_{hu} > \gamma_{ls} = \gamma_{lu}$ during regular times, but when there is an acceleration in skill bias, the pattern changes to favor workers with more unobserved skills, i.e., $\gamma_{hs} > \gamma_{hu}$ and $\gamma_{ls} > \gamma_{lu}$. Although this assumption can generate stable residual inequality before 1970s, and an increase in residual inequality during the 1970s, it is largely ad hoc—it is *reverse engineered* to fit the facts.

exogenous skill premium $\overline{\omega} = w_H/w_L$. This implies:

$$\frac{H}{l} = \left(\frac{A_h}{A_l}\right)^{\rho/(1-\rho)} \overline{\omega}^{-1/(1-\rho)}. \tag{30}$$

where the level of employment of unskilled workers, l, will generally be less than their labor supply L because of the wage compression. A more compressed wage structure—i.e., a lower $\overline{\omega}$— will therefore increase the unemployment of unskilled workers, given by $L - l$.

The view that wages are more compressed in Europe clearly has some merit. Blau and Kahn (1995) show that the major difference in overall inequality between the U.S. and many continental European economies is not in the 90-50 differential, but in the 50-10 differential. This suggests that the minimum wage, strong unions, and generous transfer programs in Europe are in part responsible for the relative wage compression in Europe.

Nevertheless, the Krugman hypothesis runs into two difficulties. First, unless there are extremely rigid institutions that fix the skill premium exogenously, skill-biased technical change should increase wage inequality irrespective of the degree of exogenously imposed wage compression. In contrast, in many continental European economies, most notably in Germany, wage inequality was very stable (see, e.g., Freeman and Katz, 1995).

Second, the Krugman hypothesis makes an explicit prediction: to the extent that wage compression is preventing the increase in the inequality of wages, profit maximizing employment decisions of firms should lead to a large decline in the employment of unskilled workers relative to that of skilled workers. In fact, skill-biased technical change might even reduce the unemployment rates of skilled workers. Yet, in Europe, the unemployment of skilled and unskilled workers increased together (e.g. Nickell and Bell, 1996, Krueger and Pischke, 1997), and unskilled employment did not grow faster in the U.S. than in European economies (Card, Kramartz and Lemieux, 1996, Krueger and Pischke, 1997).

It is possible that bargaining arrangements in Europe between firms and unions, imply not only wage compression, but deviations from the relative demand curve for skills given by (30). This can be because European institutions may be forcing firms to pay uniform wages to all educated workers irrespective of their exact contribution, making the employment of skilled workers less profitable as well. Alternatively, if unions represent both skilled and unskilled workers, and are committed to wage compression, they may not want to suffer a large decrease in the employment of unskilled workers. So they may be willing to make certain concessions in wage levels in order to induce firms to employ more unskilled workers at a compressed wage structure. Although such deviations from equation (30) are a possibility, we have no direct evidence to assess how far off the relative demand curve European economies may be, and how they would respond to skill-biased technical change in this situation. It is also useful to bear in mind that European economies, as the U.S., are likely to have experienced skill-biased technical change not only during the past thirty years, but for much longer. So how continental European economies responded to the more recent wave ofskill-biased technologies cannot be analyzed in a vacuum.

An alternative view suggested by Nickell and Bell (1996) explains the differences in the wage structure across countries by differences in the skill distribution. According to this view, because of the relative weakness of the U.S. high school system, American noncollege workers are less skilled than their European counterparts. However, recent work by Devroye and Freeman (2000) shows that differences in skill distribution have little to do with cross-country differences in wage dispersion. They document that dispersion of internationally comparable test scores among native born Americans are very similar to those in Europe, but wage inequality among native born Americans is much higher. Moreover, the Nickell-Bell approach also fails to explain the differential changes in inequality: the U.S. was roughly as unequal as France in the 1970s, and the relative test scores of American youth have not deteriorated since then.

My preferred approach to explaining cross-country differences is to consider the effect of labor market institutions on technology choices. In particular, the European labor market institutions, which compress the structure of wages, will give greater incentives to adopting labor-complementary technologies, and will reinforce wage compression. I give a simple example to illustrate the point here. Suppose the productivity of a skilled worker is $A_h = a\eta$, whereas the productivity of an unskilled worker is $A_l = a$, where a is a measure of aggregate technology in use, and $\eta > 1$. Suppose that wages are determined by rent sharing, unless they fall below a legally mandated minimum wage, in which case the minimum wage binds. Hence, $w_j = \min\{\beta A_j, \underline{w}\}$, where $j = l$ or h, and β is worker's share in rent sharing. Note that the cost of upgrading technology does not featuring in this wage equation, because rent sharing happens after technology costs are sunk. To capture wage compression, suppose the minimum wage is binding for unskilled workers in Europe. Now consider technology adoption decisions. In particular, firms can either produce with some existing technology, a, or upgrade to a superior technology, $a' = a + \alpha$, at cost γ. The profit to upgrading the technology used by a skilled worker is $(1-\beta)\alpha\eta - \gamma$, both in the U.S. and Europe. The new technology will therefore be adopted as long as

$$\gamma \leq \gamma^S \equiv (1-\beta)\alpha\eta.$$

Note that there is a holdup problem, discouraging upgrading: a fraction β of the productivity increase accrues to the worker due to rent sharing (Grout, 1984, Acemoglu, 1996).

The incentives to upgrade the technology used by unskilled workers differ between the U.S. and Europe. In the U.S., this profit is given by $(1-\beta)\alpha - \gamma$. So, the new technology will be adopted with unskilled workers if

$$\gamma \leq \gamma^U \equiv (1-\beta)\alpha.$$

Clearly, $\gamma^U < \gamma^S$, so adopting new technologies with skilled workers is more profitable. In contrast, the return to introducing the new technology is different in Europe because minimum wages are binding for unskilled workers. To simplify the discussion, suppose that even after the introduction of new technology, the minimum wage binds, i.e., $\underline{w} > \beta(A + \alpha)$. Then, the return to introducing the new technology in Europe with unskilled workers is $\alpha - \gamma$, and firms will do so as long as $\gamma < \alpha$. Since $\alpha > \gamma^U$, firms in Europe have greater incentives to introduce advanced technologies with unskilled workers than in the U.S.. Intuitively, the binding minimum wage in Europe makes the firm *the full residual claimant* of the increase in productivity of unskilled workers. This highlights that in an economy with a compressed wage structure, firms may have a greater incentive to increase the productivity of unskilled workers (see Acemoglu and Pischke, 1999, for this argument in the context of training).

As long as the cost of upgrading technologies, γ, is small, i.e., less than γ^U, new technology will be used both with skilled and unskilled workers in the U.S. and Europe. In this case, cross-country inequality levels will be stable. This corresponds to the situation in the 1950s and 1960s. In contrast, if γ is high, for example, because the technological improvements of the 1980s are more expensive to implement, there may be a divergence in inequality between the two economies. For instance, if $\gamma \in (\gamma^U, \alpha)$, then new technology will not be adopted with unskilled workers in the U.S., but it will be used with unskilled workers in Europe. As a result, while wage inequality increases in the U.S., it will remain stable in Europe. Therefore, a simple story for cross-country differences in inequality trends emerges from this model: wage compression encourages the use of more advanced technologies with unskilled workers, and acts to reinforce itself in Europe. In contrast, technological developments can harm the earnings of low-skill U.S workers who are not protected by this type of compression. Whether the interaction between wage compression and technology choice could be important in explaining European inequality and unemployment patterns is an area for future study.

8.2 International determinants of technology

The endogenous technology framework developed above links the skill bias of technology to the relative supply of skills. There are a number of interesting and difficult issues that arise when we consider the international dimension. Here I simply mention some preliminary approaches, but clearly much theoretical and empirical work remains to be done.

A first extension of the endogenous technology idea to an international context might be to suppose that skill bias in each country is determined by the country's relative supply of skills. However, there are reasons to expect that new technologies will spread across countries. In this case, it may be the incentives in the technologically most advanced country (the technological leader) that determine the skill bias of world technologies. This description may be adequate for understanding the skill bias of technologies used by less developed countries (see for example Acemoglu, 1999b). It is also possible for other technologically advanced economies to pursue a different path of technological development than the leader, in which case domestic incentives may be important in shaping skill bias.

What determines the skill bias of technologies developed by the technological leader? This depends on the market sizes for different types of technologies, hence on the international enforcement of intellectual property rights. For example, in the discussion on the effect of trade on technology, I supposed that there were no intellectual property rights for U.S. companies enforced in less developed economies. In this case, incentives to develop new technologies are shaped by the U.S. domestic supplies. This may be a good starting point, since even when property rights are enforced, there will be a number of difficulties facing U.S. companies marketing their technologies in other countries. For example, technologies may need to be adapted to conditions in local markets, or producers in LDCs may be unable to pay for these technologies because of credit problems.

It is also worth noting that even when a country is using U.S. technologies, its effective skill bias may be influenced by its domestic skill supply. This is because U.S. technologies need to be adapted to local conditions, and firms will have a greater incentive to do this when there is a larger supply of workers to use these technologies. So it may be not only technological change that is endogenous to relative supplies, but also technology adoption.

Finally, another interesting cross-country dimension comes from looking at wage inequality trends in LDCs. As discussed in Section 6.3, the first order predictions of the standard trade theory are not borne out: instead of a decline in inequality, which would have been expected due to the greater integration of these economies into world trade, inequality increased in most LDCs. The recent paper by Berman and Machin (2000) shows an interesting pattern: while there has been rapid skill upgrading in many middle income countries, there is much less evidence for rapid skill upgrading in the poorest economies. A possible explanation for these patterns is that middle income countries are adopting advanced technologies much more rapidly than the poorest countries, and since these technologies are more skill-biased, these economies are undergoing rapid skill upgrading and increases in inequality. Furthermore, if, as claimed by Acemoglu and Zillibotti (1999), new technologies developed in the rich economies are typically "too skill-biased" for LDCs, the recent acceleration in skill bias could have negative implications for the LDCs. More generally, the impact of technologies developed in the advanced economies on LDC labor markets is an area that requires further research.

9 Appendix

9.1 Data sources

The samples are constructed as in Katz and Autor (2000). I thank David Autor for providing me with data from this study. Data from 1939, 1949 and 1959 come from 1940, 1950 and 1960 censuses. The rest of the data come from 1964-1997 March CPSs. The college premium is the coefficient on workers with a college degree or more relative to high school graduates in a log weekly wage regression. The regression also includes dummies for other education categories, a quartic in experience, three region dummies, a nonwhite dummy, a female dummy, and interactions between the female dummy and the nonwhite dummy and the experience controls. The sample includes all full-time full-year workers between the ages of 18 and 65, and except those with the lowest 1 percent earnings. Earnings for top coded observations are calculated as the value of the top code times 1.5. The relative supply of skills is calculated from a sample that includes all workers between the ages of 18 and 65. It is defined as the ratio of college equivalents to non-college equivalents, calculated as an Autor, Katz and Krueger (1998) using weeks worked as weights. In particular, college equivalents=college graduates+0.5×workers with some college, and noncollege equivalents=high school dropouts+high school graduates+0.5×workers with some college.

Samples used for overall and residual wage inequality include only white male full-time full year workers between the ages of 18 and 65, and excludes those earnings less than half the real value of the 1982 minimum wage converted from nominal dollars using the personal consumption expenditure deflator (see Katz and Autor, 2000). Earnings for top coded observations are calculated as the value of the top code times 1.5.

9.2 The behavior of overall inequality during the 1970s

In an important paper on the effect of labor market institutions on inequality, DiNardo, Fortin and Lemieux (1995) provide evidence suggesting that in the May CPSs, there is no increase in inequality during the 1970s. In Table 1, I display numbers from the survey by Katz and Autor (2000), who report changes in residual inequality for the past four decades from three different sources; decennial censuses, and March CPSs and May CPSs (and later Outgoing Rotation Group files–ORGs). These numbers show no significant change in residual or overall inequality during the 1960s, and consistent increases in inequality from all sources during the 1970s and the 1980s. For example, the data from the Census and the March CPSs indicate that the 90-10 differential increased about .01 a year between 1970 and 1979, while the 90-50 differential increased by about .011 a year during the same period. The May CPS data show a smaller increase in the 90-10 differentials during this period, but a comparable increase in the 50-10 differential. Overall, although there is less uniformity among data sources regarding the behavior of residual inequality than returns to schooling (see Katz and Autor, 2000), there is considerable evidence that residual and overall inequality started to increase during the 1970s.

9.3 Can composition effects explain inequality changes?

A possible explanation for the patterns we observe could be changes in the distribution of unobserved skills—or more concretely, *composition effects*. For example, the average ability of workers with high education may have increased relative to that of workers with low education over time. Here, I document that the increase in the returns to education and residual inequality are not simply due to composition effects. Note first that composition effects cannot by themselves explain the recent changes in inequality: as noted in subsection 7.2, composition effects suggest that inequality among educated and uneducated workers should move in opposite directions. This suggests that changes in the true returns to skills have played at least some role in the changes in inequality.

More generally, to get a sense of how important composition effects may be, consider a variant of equation (13) above with two education levels, high $h = 1$ and low $h = 0$, and suppose wages are given by

$$\ln w_{it} = a_i + \gamma_t h_i + \varepsilon_{it} \tag{31}$$

where h_i is a dummy for high education, a_i is unobserved ability, and ε_{it} is a mean zero disturbance term. Define the (log) education premium—the difference between the average wages of high and low education workers—can be written as

$$\ln \omega_t \equiv E\left(\ln w_{it} \mid h_i = 0\right) - E\left(\ln w_{it} \mid h_i = 0\right) = \gamma_t + A_{1t} - A_{0t}$$

where $A_{1t} \equiv E\left(a_i \mid h_i = 0\right)$ and A_{0t} is defined similarly. The increase in the education premium can be caused by an increase in γ_t (a true increase in the returns to skills) or an increase in $A_{1t} - A_{0t}$. There are basically two reasons for an increase in $A_{1t} - A_{0t}$: (1) changes in cohort quality, or (2) changes in the pattern of selection into education.

Consider changes in cohort quality first. If, as many claim, the U.S. high school system has become worse, we might expect a decline in A_{0t} without a corresponding decline in A_{1t}. As a result, $A_{1t} - A_{0t}$ may increase. Alternatively, as a larger fraction of the U.S. population obtains higher education, it is natural that selection into education (i.e., the abilities those obtaining education) will change. It is in fact possible that those who are left without education could have very low unobserved ability, which would translate into a low level of A_{0t}, and therefore into an increase in $A_{1t} - A_{0t}$.

Although these scenarios are plausible, theoretically the opposite can happen as well. For example, many academics who have been involved in the U.S. education system for a long time complain about the decline in the quality of universities, while the view that American high schools have become much worse is not shared universally (e.g., Krueger, 1998). The selection argument is also more complicated than it first appears. It is true that, as long as those with high unobserved abilities are more likely to obtain higher education, an increase in education will depress A_{0t}. But it will also depress A_{1t}. To see why assume that there is perfect sorting—i.e., if an individual with ability a obtains education, all individuals with ability $a' > a$ will do so as well. In this case, there will exist a threshold level of ability, $\overline{a}$, such that only those with $a > \overline{a}$ obtain education. Next consider a uniform distribution of a_i between b_0 and $b_0 + b_1$. Then, $A_0 = \frac{1}{\overline{a} - b_0} \int_{b_0}^{\overline{a}} a da = \frac{\overline{a} + b_0}{2}$ and $A_1 = \frac{1}{b_1 - b_0 - \overline{a}} \int_{\overline{a}}^{b_0 + b_1} a da = \frac{b_0 + b_1 + \overline{a}}{2}$. So both A_0 and A_1 will decline when $\overline{a}$ decreases to $\overline{a}'$. Moreover, $A_1 - A_0 = b_1/2$, so it is unaffected by the decline in $\overline{a}$. Intuitively, with a uniform distribution of a_i, when $\overline{a}$ increases, both A_0 and A_1 fall by exactly the same amount, so the composition effects have no influence on the education premium. Clearly, with other distributions of ability, this extreme result will no longer hold, but it remains true that both A_0 and A_1 will fall, and whether this effect will increase or decrease the education premium is unclear. Overall, therefore, the effects of changes in composition on education premia is an empirical question.

Empirically, the importance of composition effects can be uncovered by looking at inequality changes by cohort (see Blackburn, Bloom and Freeman, 1992; Juhn, Murphy and Pierce,1993). To see this, rewrite equation (31) as

$$\ln w_{ict} = a_{ic} + \gamma_t h_{ic} + \varepsilon_{cit} \tag{32}$$

where c denotes a cohort—i.e., a group of individuals who are born in the same year. I have imposed an important assumption in writing equation (32): returns to skills are assumed to be the same for all cohorts and ages; γ_t—though clearly they vary over time. We can now define cohort specific education premia as

$$\ln \omega_{ct} \equiv E\left(\ln w_{ict} \mid h_i = 0\right) - E\left(\ln w_{ict} \mid h_i = 0\right) = \gamma_t + A_{1ct} - A_{0ct}$$

where $A_{1ct} \equiv E\left(a_{ic} \mid h_i = 0\right)$ and A_{0ct} is defined similarly. Under the additional assumption that there is no further schooling for any of the cohorts over the periods under study, we have $\ln \omega_{ct} = \gamma_t + A_{1c} - A_{0c}$, which implies

$$\Delta \ln \omega_{c,t'-t} \equiv \ln \omega_{ct'} - \ln \omega_{ct} = \gamma_{t'} - \gamma_t, \tag{33}$$

i.e., changes in the returns to education within a cohort will reveal the true change in the returns. The assumption that returns to skills are constant over the lifetime of an individual may be too restrictive. Murphy and Welch (1992), for example, show quite different age earning profiles by education. Nevertheless, a similar argument can be applied in this case too. For example, suppose$\ln \omega_{cst} = \gamma_{st} + A_{1c} - A_{0c}$ for cohort c of age s in year t, and that $\gamma_{st} = \gamma_s + \gamma_t$ (this assumption is also not necessary, but simplifies the discussion). Then $\Delta \ln \omega_{c,t'-t} = \gamma_{s'} - \gamma_s + \gamma_{t'} - \gamma_t$, where obviously $s' - s = t' - t$. Now consider a different cohort, c'' that is age s' in the year t and age s in the year t''. Then $\Delta \ln \omega_{c'',t-t''} = \gamma_{s'} - \gamma_s + \gamma_t - \gamma_{t''}$. So, the true change in the returns to skills between the dates t'' and t' is

$$\Delta^2 \ln \omega \equiv \Delta \ln \omega_{c,t'-t} - \Delta \ln \omega_{c'',t-t''} = \gamma_{t'} - \gamma_{t''}. \tag{34}$$

Using data from the 1950-1990 censuses, Table A2 gives some of the single and double differences of cohort inequality for white men aged 26-55. The single differences show increases in the returns to college within most cohorts, with the exception of the years between 1970 and 1980. Therefore, these increases are likely to reflect differential age effects by education. In contrast, the numbers in Panel C for the 1950-70 period show no increases, suggesting that the double difference does a good job of controlling for composition effects. The numbers for the 1960-80 period are negative, which likely reflect the decline in the college premium between 1960 and 1980. The final row gives the most important results of this table. The 1970-90 double differences are large and positive, suggesting that the true returns to education increased over this time period. Interestingly, despite the well-known evidence that the college premium increased faster for younger workers over the 1980s, the results in Table A2 show that the true increase in returns to skills between 1970 and 1990 are comparable for cohorts born between 1936 and 1955. These results therefore indicate that the major component of the increase in that college premium during the 1980s and 90s was changes in skill prices, not composition effects.

Table A3, which replicates Table 3 from Juhn, Murphy and Pierce (1993), shows that the increase in overall and residual inequality cannot be explained by composition effects either. Panel A shows that the 90-10 differential for cohorts entering the market between 1935 and 1964 is approximately constant between 1964 and 1970, but increases sharply for each cohort between 1970 and 1976, and then increases further in 1982 and 1988. Panel B shows a similar picture for log wage residuals. These results suggest that the changes in the structure of wages observed over the past 30 years cannot be explained by pure composition effects, and reflect mainly changes in the true returns to observed and unobserved skills.

References

Acemoglu, Daron, “A Microfoundation for Social Increasing Returns in Human Capital,” *Quarterly Journal of Economics*, 111 (1996), 779–804.

Acemoglu, Daron, “Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality” *Quarterly Journal of Economics*, 113 (1998), 1055-1090.

Acemoglu, Daron “Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence” *American Economic Review*, 89(1999a), 1259-1278.

Acemoglu, Daron, “Patterns of Skill Premia”, NBER Working Paper No. 7018, (1999b).

Acemoglu, Daron, “Labor- and Capital-Augmenting Technical Change”, NBER Working Paper No. 7544, (2000).

Acemoglu, Daron, Phillipe Aghion and Gianluca Violante “Technical Change, Deunionization and Inequality MIT mimeo, (2000).

Acemoglu, Daron and Jörn-Steffen Pischke “The Structure of Wages an Investment in General Training” *Journal of Political Economy*, 107, (1999), 539-572.

Acemoglu, Daron and Fabrizio Zilibotti, “Productivity Differences” (1998), forthcoming *Quarterly Journal of Economics.*

Aghion, Philippe, “Technical Change, Institutions, and the Dynamics of Inequality” mimeo, (2000).

Aghion, Philippe and Peter Howitt, *Endogenous Growth Theory*, Cambridge, MA, MIT Press, 1998.

Aghion, Philippe, Peter Howitt, and Gianluca Violante, “Technology, Knowledge, and Inequality” Working paper, 2000.

Ahmad, Syed, “On The Theory of Induced Invention,” *Economic Journal* LXXVI, (1966), pp. 344-357.

Angrist, Joshua, “The Economic Returns to Schooling in the West Bank and Gaza Strip” *American Economic Review,* 85,1065-1087.

Appelbaum Eileen and Rosemary Batt *The New American Workplace* ILR Press, New York, 1994.

Autor, David, Alan Krueger and Lawrence Katz, “Computing Inequality: Have Computers Changed the Labor Market?” *Quarterly Journal of Economics*, CXIII (1998), 1169-1214.

Autor, David, Frank Levy and Richard Murnane “Computer Cognition and Human Cognition: The Skill Content of Recent Technological Change” MIT mimeo, (2000).

Bailey, Martin and Robert Gordon “The Productivity Slowdown: Measurement Issues and Explosion of Computer Power” *Brookings Papers on Economic Activity*, macroeconomics, volume 2, (1988), 379-431.

Bairoch, Paul (1988), *Cities and Economic Development: From the Dawn of History to the Present*, University of Chicago Press, Chicago, translated by Christopher Braider.

Bartel, Ann. and Nachum Sicherman, “Technological Change and the Skill Acquisition of Young Workers,” *Journal of Labor Economics* 16, (1998), 718-755.

Bartel, Ann P. and Frank Lichtenberg, “The Comparative Advantage of Educated Workers in Implementing New Technology,” *Review of Economics and Statistics,* 49 (1987), 1-11.

Beaudry, Paul and David Green “What is Driving US and Canadian Wages: Exogenous Technical Change or Endogenous Choice of Technique?” NBER Working Paper, (1999).

Becker, Gary *Human capital.* Chicago: The University of Chicago Press, (1964).

Berman, Eli, John Bound and Zvi Griliches, “Changes in the Demand for Skilled Labor within U. S. Manufacturing Industries: Evidence from the Annual Survey of Manufactures,” *Quarterly Journal of Economics,* 109 (1994), 367-365.

Berman, Eli, John Bound and Stephen Machin “Implications of Skill-Biased Technological Change: International Evidence,” *Quarterly Journal of Economics*, CXIII (1998), 1245-1280.

Berman, Eli, and Stephen Machin, "SBTC Happens! Evidence on the Factor Bias of Technological Change in Developing and Developed Countries", mimeo, 2000.

Blackburn, David Bloom and Richard Freeman (1992) "Changes in Earnings Differentials in the 1980s: Concordance, Convergence, Causes, and Consequences" NBER Working Paper.

Blau, Francine D. and Lawrence M. Kahn "International differences in male wage inequality: Institutions versus market forces," *Journal Political Economy* 104, (1996) 791-837.

Borjas, George J., Richard B. Freeman, and Lawrence F. Katz, "How much do Immigration and Trade Affect Labor Market Outcomes?," *Brookings Papers on Economic Activity,* 1 (1997), 1-90.

Borjas, George J. and Valerie Ramey, "Foreign Competition, Market Power and Wage Inequality," *Quarterly Journal of Economics,* 110 (1995), 1075-1110.

Boskin Michael et al "Consumer Prices, The Consumer Price Index, and The Cost of Living," *Journal of Ec. Perspectives* vol. 12, (1998), 3-26.

Bound, John and George Johnson "Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations" *American Economic Review*, June 1992, 82 (3), pp 371-92.

Braverman, Harry, *Labor and Monopoly Capital: The Degradation of Work in the Twentieth Century*, New York: Monthly Review Press, 1974.

Bresnahan, Timothy F., "Computerization and Wage Dispersion: An Analytical Reinterpretation," Stanford University Mimeo (1999).

Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt, "Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-level Evidence," NBER Working Paper 7136, May 1999.

Brendt, Ernie, Catherine Morrison and Larry Rosenblum (1994) "High-tech Capital Formation and Labor Composition in U.S. Manufacturing Industries: An Exploratory Analysis" *Journal of Econometrics*, volume 65, 9-43.

Cappelli, Peter and Steffanie Wilk "Understanding Selection Processes: Organization Determinants and Performance Outcomes" Wharton School Mimeo, 1997.

Card, David E. and Alan Krueger, "School Quality and Black-White Relative Earnings: A Direct Assessment," *Quarterly Journal of Economics* 107, (1992), 151-200.

Card, David and Thomas Lemieux, "Can Falling Supply Explain the Rising Return to College for Younger Men?," University of California, Berkeley Mimeo (1999).

Card, David, Francis Kramartz and Thomas Lemieux ;"Changes in the Relative Structure of Wages and Employment: A Comparison of the United States, Canada and France" Mimeo (1996).

Caroli, Eve and John van Reenen, "Wage Inequality and Organizational Change", mimeo UCL. (1999)

Caselli, Francesco "Technological Revolutions" *American Economic Review* (1999).

Cohen, Sarit and Chang-Tai Hsieh, "Macroeconomic and Labor Market Impact of Russian Immigration In Israel" Princeton University mimeo (2000). In

David, Paul, *Technical Choice, Innovation and Economic Growth: Essays on American and British Experience in the Nineteenth Century*, London, Cambridge University Press, 1975.

David, Paul, "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox" *American Economic Review*. Vol. 80 (2), (1990) 355-61.

Davis, Steven J.; "Cross-Country Patterns of Changes in Relative Wages" *NBER Macroeconomic Annua*l, 1992 MIT Press, Cambridge, 239-292.

Davis, Steven J., John Haltiwanger, and Scott Schuh *Job Creation and Destruction*the MIT Press Cambridge, 1996.

Devroye, Dan and Richard Freeman, "Does Inequality in Skills Explain Inequality of Earnings Across Countries?" Harvard University mimeo (2000).

Diebold, Francis, David Neumark and Daniel Polskthey "Job Stability in United States" *Journal of Labor Economics*15, (1997), 206-33.

Doms, Mark, Timothy Dunne and Kenneth Torske "Workers, Wages and Technology" *Quarterly Journal of Economics*, Vol 112, (1997), 217-252

DiNardo, John, Nicole Fortin and Thomas Lemieux "Labor market institutions, and the distribution of wages, 1973-1992: A semiparametric approach," *Econometrica* 64, (1995), 1001-1044.

DiNardo, John and Jörn-Steffen Pischke "The returns to computer use revisited: Have pencils changed the wage structure too?" *Quarterly Journal of Economics* 112, (1997), 291-303.

Farber, Henry, "Are Lifetime Jobs Disappearing? Job Duration in United States" in the are working paper, No. 5014.

Foster, Andrew and Mark Rosenzweig "Technical Change in Human Capital Return and Investments:Evidence from the Green Revolution" *American Economic Review*, 86, (1996), 931-953.

Frank, Robert H. and Phillip J. Cook, "The Winner-Take-All Society," *Free Press*, New York, (1995).

Freeman, Richard B., "Overinvestment in College Training?," *Journal of Human Resources,* 10 (1975), 287-311.

Freeman, Richard, *The Overeducated American*, New York, Academic Press, 1976.

Freeman, Richard "Demand For Education" Chapter 6 in Orley Ashenfelter and Richard Layard (editors) *Handbook of Labor Economics*, North Holland, Vol I (1986), 357-386.

Freeman, Richard, "How Much Has De-unionization Contributed to the Rise of Male Earnings Inequality?", NBER working paper 3826, 1991

Freeman, Richard and Lawrence F. Katz, "Introduction and Summary," in R. Freeman and L. Katz (eds.), *Differences and Changes in Wage Structures,* Chicago, IL: The University of Chicago Press, pp. 1-22

Freeman Richard and Edward Lazear "An Economic Analysis of Works Councils ", in Rogers,-Joel;Streeck,-Wolfgang, eds. *Works councils: Consultation, representation, and cooperation in industrial relations.* National Bureau of Economic Research Comparative Labor Markets Series. Chicago and London: University of Chicago Press, 1995, pages 27-50, (1995).

Freeman, Richard and James Medoff, *What Unions Do*, Basic Books, New York, 1984.

Freeman, Richard and Karen Needles, "Skill Differences in Canada in an Era of Rising Labor Market Inequality," in D. Card and R. Freeman, eds., *Small Differences That Matter*, University of Chicago Press and NBER, Chicago, IL (1993).

Friedberg, Rachel, "The Impact of of Mass Migration on the Israeli Labor Market," Brown University Population Studies and Training Center Working Paper #97-11, September 1997.

Galor Oded and Daniel Tsiddon, "Technological Progress, Mobility and Economic Growth" *American Economic Review*, LXXXVII, pp 363-382.

Galor, Oded and Omer Maov "Ability Biased Technological Transition, Wage Inequality and Economic Growth" *Quarterly Journal of Economics* 115 (2000).

Gandal, Neil and Gordon H. Hanson, and Matthew J.Slaughter, "Technology, Trade, and Adjustment to Immigration in Israel," Dartmouth College mimeo, 2000.

Goldin, Claudia and Robert Margo, "The Great Compression: The Wage Structure in the United States at Mid-Century," *Quarterly Journal of Economics* 107 (1992), 1-34.

Goldin, Claudia and Lawrence Katz, "The Decline of Non-Competing Groups: Changes in the Premium to Education, 1890 to 1940" NBER W.P. No. 5202, 1995.

Goldin, Claudia and Lawrence F. Katz, "The Origins of Technology-Skill Complementarity," *Quarterly Journal of Economics* 113 (1998), 693-732.

Goldin, Claudia and Lawrence Katz, "Education and Income in the Early 20th-century: Evidence from the Prairies" 2000, forthcoming *Journal of Economic History.*

Gosling, Amanda, "The Determination of Union Status and Wages in British Establishments", Discussion paper 491, University of Essex, 1998.

Gottschalk, Peter, "Inequality in Income, Growth and Mobility: The Basic Facts," *Journal of Economic Perspectives* 11 (1997), 21-40.

Gottschalk, Peter and Robert Moffitt, "The Growth of Earnings Instability in the U. S. Labor Market," *Brookings Papers on Economic Activity* 2 (1994), 217-272.

Gottschalk, Peter and Timothy M. Smeeding, "Cross-National Comparisons of Earnings and Income Inequality," *Journal of Economic Literature* 35 (1997), 633-687.

Gordon, Robert *Measurement of Durable Goods Prices,* University of Chicago Press, 1990.

Gordon, Robert "Current Productivity Puzzles from a Long-term Perspective" Northwestern University mimeo, (1998).

Gould, Eric D., Omer Moav, and Bruce A. Weinberg, "Precautionary Demand for Education, Inequality, and Technological Progress," Mimeo, Feb 2000.

Green, Constance M., *Eli Whitney and The Birth of American Technology*, Boston, Little, Brown (1956).

Greenwood, Jeremy , Zvi Hecowtiz and Per Krusell, "Long-Run Implications of Investment-Specific Techological Change," *American Economic Review,* vol. 87 [3], (1997) 342-362.

Greenwood, Jeremy and Mehmet Yorukoglu, "1974" *Carnegie-Rochester Conference Series on Public Policy*, XLVI (1997), 49-95.

Griliches, Zvi, "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, Vol. 25 (1957) 501-522.

Griliches, Zvi, "Capital-Skill Complementarity," *Review of Economics and Statistics* 51 (1969), 465-468.

Grout, Paul, "Investment and Wages in the Absence of Binding Contracts: A Nash Bargaining Approach," *Econometrica*, 52 (1984), 449–460.

Habakkuk, H.J., *American and British Technology in the Nineteenth Century*, London: Cambridge University Press, 1962.

Hanson, Gordon H. and Ann Harrison, "Trade Technology, and Wage Inequality: Evidence from Mexico," (September 1994), mimeo.

Hayami, Yujiro and Vernon Ruttan, "Factor Prices and Technical Change in Agricultural Development: the U.S. and Japan, 1880-1960" *Journal of Political Economy*, LXXII (1970), 1115-1141.

Hornstein, Andres and Per Krusell (1996), "Can Technology Improvements Cause Productivity Slowdowns?," *NBER Macroeconomic Annual*, vol. 11, pp. 209-259.

Ichinowski, Casey, Kathryn Shaw and Giovanna Prennushi (1997);"The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines" *American Economic Review* Vol 87, pp 291-313.

James, John A., and Jonathan S. Skinner, (1985) "The Resolution of the Labor-Scarcity Paradox," *Journal of Economic History*, Volume 45, pp.513 - 540.

Jerome, Harry, *Mechanization in History*, New York: National Bureau of Economic Research, 1934.

Johnson, George "Changes in Earnings Inequality: The Role of Demand Shifts" *Journal of Economic Perspectives*, 11 (1997), 41-54.

Johnson, George and Frank Stafford, "The Labor Market Implications of International Trade" *Handbook of Labor Economics* volume 3.

Jorgensen, Dale, W. and Kevin J. Stiroh, "Raising the Speed Limit: U.S. Economic Growth in the Information Age" forthcoming in the *Brookings Papers on Economic Activity* (2000).

Juhn, Chinhui, Kevin M. Murphy and Brooks Pierce (1993);"Wage Inequality and the Rise in Returns to Skill" *Journal of Political Economy* Vol 101, pp 410-442.

Katz, Lawrence, and David Autor, "Changes in the Wage Structure and Earnings Inequality" in O. Ashenfelter and D. Card, eds., *The Handbook of Labor Economics Volume III*, Amsterdam: Elsevier.

Katz, Lawrence and Kevin Murphy, "Changes in Relative Wages: Supply and Demand Factors" *Quarterly Journal of Economics*, CVII (1992), 35-78.

Katz, Lawrence, Gary W. Loveman and David G. Blanchflower "A Comparison of Changes in the Structure of Wages in Four OECD Countries" in Richard Freeman and Lawrence Katz op. cit.

Kendrick, John W, *Productivity Trends in the United States,* Princeton [N.J.] Princeton University Press, 1961.

Kiley, Michael, "The Supply of Skilled Labor and Skill-Biased Technological Progress" 1999, *Economc Journal.*

Kremer, Michael and Eric Maskin, "Segregation By Skill and the Rise in Inequality" MIT and Harvard Mimeo, 1999.

Krueger, Alan, "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989" *Quarterly Journal of Economics*, CIIX (1993), 33-60.

Krueger, Alan, "Labor Market Shifts and the Price Puzzle Revisited" Princeton University, Industrial Relations Section Working Paper. No. 375, 1997.

Krueger, Alan, "Reassessing the View That American Schools Are Broken" Industry Relation Section Princeton University Working Paper 1998.

Krueger, Alan, "Measuring Labor's Share" *American Economic Review*, 89 (1999), 45-51.

Krueger, Alan B., and Jörn-Steffen Pischke "Observations and conjectures on the U.S. employment miracle," NBER Working Paper No. 6146, (1997).

Krugman, Paul, "Part and Prospective Causes of High Unemployment" in Federal Reserve Bank of Kansas City *Reducing Unemployment: Current Issues and Policy Options* (1994)

Krugman, Paul, "Growing World Trade: Causes and Consequences" *Brookings Papers on Economic Activity* (1995), pp 327-377..

Krusell, Per, and Lee Ohanian, and Victor Rios-Rull and Giovanni Violante, "Capital Skill Complementary and Inequality", University of Rochester Mimeo, 2000, forthcoming *Econometrica.*

Lawrence, Robert and Matthew Slaughter "Trade and U.S. Wages: Giant Sucking Sound or Small Hiccup?" *Brookings Papers on Economic Activity*, (1993) pp. 161-226.

Leamer, Edward E., "Trade, Wages and Revolving-Door Ideas," NBER Working Paper No. 4716 (1995).

Lee, David S., "Wage Inequality in the U. S. During the 1980s: Rising Dispersion or Falling Minimum Wage?," *Quarterly Journal of Economics* 114 (1999), 941-1024.

Machin, Stephen, John Van Reenen, "Technology and Changes in Skill Structure: Evidence from Seven OECD Countries," *Quarterly Journal of Economics* 113 (4), Nov. 1998, 1215-44.

Mobius, Marcus "The Evolution of Work" MIT mimeo (2000).

Mokyr, Joel, *The Levers of Riches: Technological Creativity and Economic Progress* Oxford University Press, New York, 1990.

Mokyr, Joel, ed., *The British Industrial Revolution: An Economic Perspective*, San Francisco: Westview Press, 1993.

Murnane, Richard J. and Frank Levy, *Teaching the Basic New Skills*, New York, The Free Press, 1996.

Murphy, Kevin M. and W. Graig Riddell and Paul M. Romer, "Wages, Skills and Technology in the United States and Canada," in E. Helpman, ed., *General Purpose Technologies* Cambridge, MA, MIT Press, (1998).

Murphy, Kevin M. and Finis Welch "The Structure of Wages" *Quarterly Journal of Economics*, CVII (1992), 255-285.

Murphy, Kevin M. and Finis Welch, "Industrial Change and the Rising Importance of Skill," in S. Danziger and P. Gottschalk, eds., *Uneven Tides* New York, Russell Sage, (1993)..

Nelson, Richard and Edmund Phelps "Investment in Humans, Technological Diffusion and Economic Growth" *American Economic Association Papers and Proceedings*, volume 56, 69-75, (1966).

Newell, Richard, Adam Jaffee and Robert Stavins, "The Induce Innovation Hypothesis and Energy-Saving Technological Change" *Quarterly Journal of Economics*, CXIV (1999), 907-940.

Nickell Stephen and Brian Bell "The Collapse in Demand For the Unskilled and Unemployment Across the OECD" *Oxford Review of Economic Policy* Vol 11, (1996) pp 40-62.

Neumark, David, Daniel Polsky, and Daniel Hansen "Has Job Stability Declined Yet? Evidence for the 1990s" *Journal of Labor Economics*, 17, (1999), S29-65.

Nordhaus William D. "Quality Change in Price Indexes," *Journal of Economic Perspectives,* 12, (1998), 59-68.

Nye, David E., *Electrifying America: Social Meanings of a New Technology, 1880-1940*, Cambridge: The Massachusetts Institute of Technology Press, 1990.

OECD *Jobs Study* Volumes 1 and 2 (1994).

Oliner, Stephen and Daniel Sichel, "Computers and Output Growth Revisited: How Big Is the Puzzle?" *Brookings Papers on Economic Activity*, (1994) pp. 273-318.

Piore, Michael and Charles Sabel, *The Second Industrial Divide,* Basic Books, New York, 1984.

Reynolds, Lloyd G. *The Structure of Labor Markets; Wages and Labor Mobility in Theory and Practice,* New York, Harper (1951).

Robbins, Donald J., "Trade and Inequality in Seven Developing Countries: Summary and Synthesis of Country Studies," (December 1995), mimeo.

Rubinstein, Yonna and Dani Tsiddon, "Coping with Technological Progress: The Role of Ability in Making Inequality so Persistent," Tel-Aviv University working paper, (1998).

Sachs, Jeffrey and Howard Shatz "Trade and Jobs in US Manufacuring" *Brookings Papers on Economic Activity*, (1994) pp. 1-84.

Schiff, Ze'ev and Ehud Ya'ari, *Intifada, The Palestinian Uprising - Israel's Third Front*, New York: Simon and Schuster, 1989.

Schmookler, Jacob, *Invention and Economic Growth*, Cambridge, Harvard University Press, 1966.

Schultz, Theodore, *The Economic Value Education*, Columbia University Press, New York (1967).

Schultz, Theodore, "The Value of the Ability to Deal with Disequalibria," *Jouranl of Economic Literature* 13, (1975) 827-846.

Thesmar, David and Mathias Thoning "Creative Destruction and Frim Organization Choice: A New Look into the Growth-Inequality Relationship" forthcoming *Quarterly Journal of Economics*, (2000).

Tinbergen, Jan, *Income Difference: Recent Research*, Amsterdam: North-Holland Publishing Company, 1975.

Violante, Giovanni L. "Technological Acceleration, Skill Transferability and the Rise in Residual Inequality", mimeo, University College London (1999).

Welch, Finis, "Education in Production," *Journal of Political Economy* 78 (1970), 312-327.

Welch, Finis, "Black-White Differences in Returns to Schooling," *American Economic Review* 63, (1973) 893-907.

Williamson, Jeffrey, *Coping With City Growth During the British Industrial Revolution*, Cambridge, Cambridge University Press, 1990.

Williamson, Jeffrey G. and Peter H. Lindert, *American Inequality: A macroeconomic History*, New York: Academic Press, 1980.

Wood, Adrian, *North-South Trade, Employment and Inequality: Changing Fortunes in a Skill Driven World*, Oxford, Clarendon Press, 1994.

Woolf, Arthur G., "Electricity, Productivity, and Labor Saving: American Manufacturing, 1900 - 1929," *Explorations in Economic History*, Vol. 21, (1984), 176-191.

Zuboff, Shohana *In the Age of the Smart Machine: The Future of Work and Power* Basic Books, New York (1988).

Table 1: Employment Shares and Skill-Biased Technical Change 1940-1990

	Employment share			Wage Bill Share		
	Some col.	Col. grad	Col. equi.	Some col.	Col. grad.	Col. equi
1940	6.4	6.1	9.3	8.9	12.3	16.7
1950	9.5	7.7	12.4	11.0	11.9	17.4
1960	12.5	10.1	16.4	14.1	16.4	23.4
1970	16.4	13.4	21.5	16.5	21.5	29.7
1980	23.6	19.2	31.0	22.4	28.1	39.3
1990	30.8	24.0	39.3	28.5	36.7	51.0

	$\sigma = 1.4$						$\sigma = 2$					
	Some col.		Col. grad		Col. equi.		Some col.		Col. grad		Col. equi	
	$\frac{A_h}{A_l}$	D	$\frac{A_h}{A_l}$	D	$\frac{A_h}{A_l}$	D	$\frac{A_h}{A_l}$	D	$\frac{A_h}{A_l}$	D	$\frac{A_h}{A_l}$	D
1940	0.004	0.21	0.016	0.31	0.035	0.38	0.140	0.37	0.303	.055	0.392	0.63
1950	0.006	0.24	0.011	0.28	0.030	0.37	0.146	0.38	0.219	0.47	0.313	0.56
1960	0.013	0.29	0.030	0.37	0.080	0.48	0.189	0.43	0.343	0.59	0.476	0.69
1970	0.017	0.32	0.069	0.47	0.179	0.61	0.199	0.45	0.485	0.70	0.652	0.81
1980	0.042	0.40	0.157	0.59	0.486	0.81	0.270	0.52	0.643	0.80	0.933	0.97
1990	0.090	0.50	0.470	0.81	1.777	1.18	0.357	0.60	1.064	1.03	1.673	1.29

Note: The first panel gives the ratio of the employment of skilled relative to unskilled, and the wage bill of skilled to unskilled workers for the corresponding skill categories. These data are taken from Autor, Katz and Krueger (1998). Some college refers to those with more than a high school (hence the measure is those with more than high school divided by those with high school or less). College graduate refers to all of those with a college degree, and college equivalent is defined as in Autor et al. It is those with a college degree+ 0.5 $\times$those with some college (correspondingly, the unskilled are defined as those with high school and less +0.5 $\times$those with some college). The bottom panel gives the implied technology shifts using equations (8) and (9) above for different values of the elasticity of substitution. The demand index D is defined as $(A_h/A_l)^{\frac{\sigma-1}{\sigma}}$.

Table 2: The effect of the relative price of equipment on skilled premia

Dependent variable is log college premium

	(1)	(2)	(3)	(4)	(5)
relative	-0.742	-0.388	-0.610	-0.691	-0.740
supply	(0.053)	(0.037)	(0.068)	(0.100)	(0.054)
time	0.026			0.022	0.024
	(0.002)			(0.007)	(0.005)
log relative price		-0.323		-0.051	
		(0.024)		(0.084)	
relative price			-0.875		-0.056
			(0.086)		(0.167)
Adjusted R^2	0.900	0.864	0.795	0.898	0.897

Note: This table reports the regression of the log college premium on a linear time trend, the log relative supply of skilled workers and various measures of the relative price of equipment capital. For comparability, all data taken from Krusell, Ohanian, Rios-Rull and Violante (2000).

Table A1–Annualized changes in overall and residual wage inequality (from Katz and Autor)

	Census		March CPSs		May CPSs-(ORGs)	
	90-10	50-10	90-10	50-10	90-10	50-10
Changes in overall inequality						
1960s	0.10	0.03	-0.03	-0.11	—	—
1970s	0.10	0.11	0.10	0.11	0.01	0.10
1980s	0.17	0.06	0.20	0.09	0.26	0.10
1990s	—	—	0.11	-0.03	0.05	0.00
Changes in residual inequality						
1960s	0.03	0.01	-0.01	-0.01	—	—
1970s	0.09	0.05	0.11	0.08	0.11	0.08
1980s	0.07	0.02	0.12	0.06	0.15	0.08
1990s	—	—	0.07	0.03	0.06	0.02

Note: The numbers give 10×annualized changes from Table 4 of Katz and Autor (2000). 90-10 is the difference between the 90th and 10th percent of the log wage or residual distribution, and 50-10 is the difference between the median and 10th percent of the corresponding distribution. The residuals are estimated from log earnings regressions with nine education dummies, a quartic in experience and their interactions. See notes to Tables 3 and 4 in Katz and Autor (2000).

Table A2: Composition Effects

Born in 19- Year↓ ⟶	06-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45	46-50	51-55
Panel A										
1950	1.448	1.370	1.175	1.093						
1960	1.551	1.564	1.525	1.421	1.303	1.132				
1970			1.680	1.656	1.613	1.539	1.392	1.153		
1980					1.567	1.560	1.538	1.402	1.222	1.063
1990							1.798	1.761	1.723	1.674
Panel B										
$\Delta \ln \omega_{50-60}$	0.103	0.194	0.350	0.328						
$\Delta \ln \omega_{60-70}$			0.155	0.234	0.311	0.407				
$\Delta \ln \omega_{70-80}$					-0.047	0.021	0.146	0.249		
$\Delta \ln \omega_{80-90}$							0.260	0.359	0.500	0.611
Panel C										
$\Delta^2 \ln \omega_{50-70}$			0.051	0.040	-0.040	0.079				
$\Delta^2 \ln \omega_{60-80}$					-0.201	-0.213	-0.165	-0.158		
$\Delta^2 \ln \omega_{70-90}$							0.307	0.338	0.354	0.362

Note: The top panel gives the college premium from the Census indicated at the beginning of the row for cohorts born in the five year intervals indicated at the head of the column. For example, the first number is for individuals born between 1906-10 from the Census of 1950. The college premium is defined as the wages of workers from that cohort with a college degree or more divided by the wages of workers from that cohort with twelve years of schooling. The bottom panel gives the change in the college premium for a given cohort between the two indicated dates and the difference between the wage growth of two neighboring cohorts as indicated by equations (33) and (34). All data are from the decennial censuses for white males born in the U.S..

Table A3: Changes in Inequality by Cohort (from Juhn et al, 1993)

Panel A: 90-10 Differentials for Log Weekly Wages					
Year of market entry	1964	1970	1976	1982	1988
1983-88					1.38
1977-82				1.27	1.38
1971-76			1.13	1.24	1.38
1965-70		1.08	1.12	1.29	1.42
1959-64	1.13	1.01	1.13	1.30	1.40
1953-58	1.02	1.07	1.16	1.32	1.43
1947-52	1.02	1.11	1.15	1.30	
1941-46	1.02	1.07	1.16		
1935-40	1.06	1.09			
1929-34	1.09				
Panel B: 90-10 Differentials for Log Wage Residuals					
Year of market entry	1964	1970	1976	1982	1988
1983-88					1.09
1977-82				1.06	1.16
1971-76			.96	1.09	1.18
1965-70		.86	.96	1.12	1.23
1959-64	.92	.86	.98	1.12	1.21
1953-58	.88	.91	.99	1.15	1.26
1947-52	.89	.94	.99	1.14	
1941-46	.94	.94	1.05		
1935-40	.95	.98			
1929-34	.99				

Note: This table replicates Table 3 of Juhn, Murphy and Pierce (1993). The top panel reports the 90-10 differential for log weekly wages of the cohorts that have entered the labor market in the corresponding six year interval. Panel B gives the 90-10 differential for the residuals from a regression of log weekly wages on education controls.

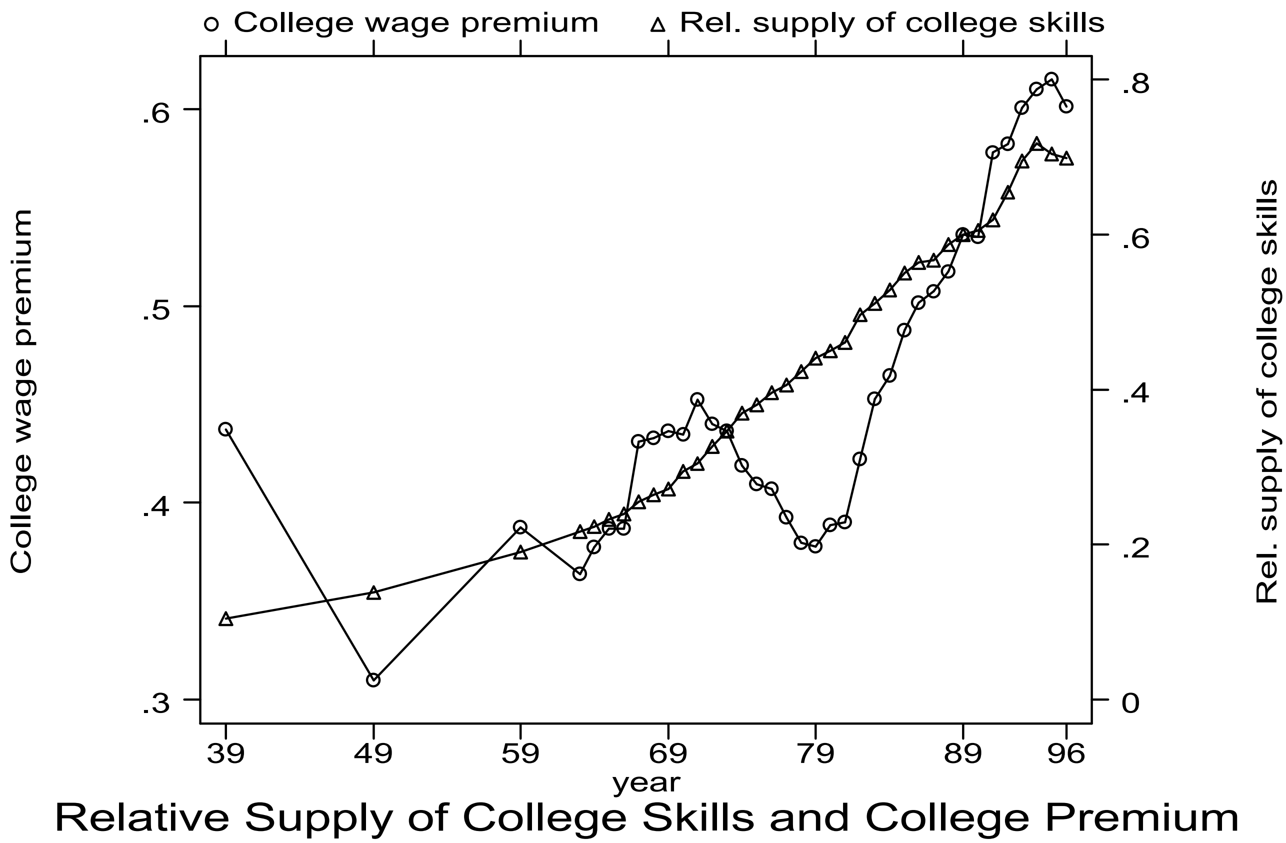

Figure 1: The behavior of the (log) college premium and relative supply of college skills (weeks worked by college equivalents divided by weeks worked of noncollege equivalents) between 1939 and 1996. Data from March CPSs and 1940, 1950 and 1960 censuses.

Figure 2: Changes in the indexed value of the 90th, 50th and 10th percentiles of the wage distribution for white males (1963 values normalized to 100).

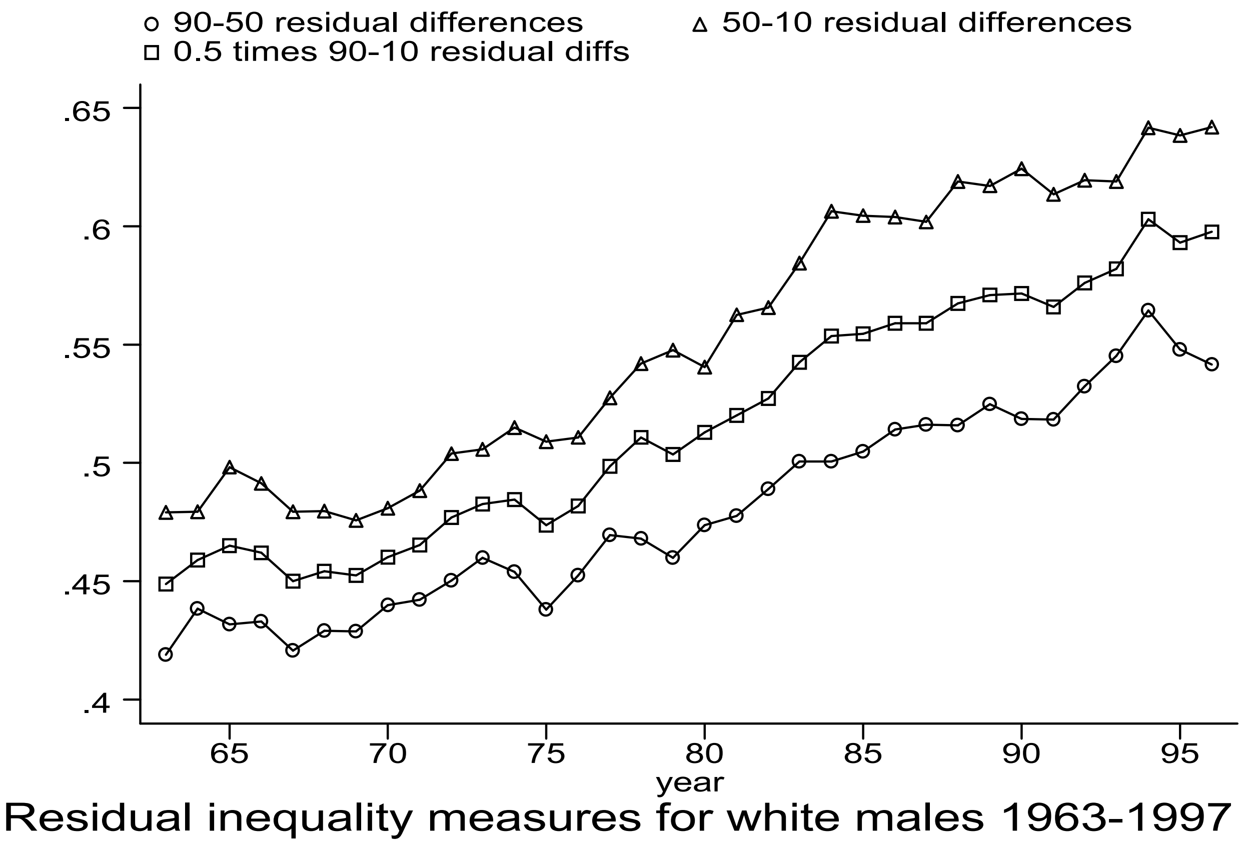

Figure 3: 90-50, 50-10 and 0.5×90-10 differentials from log weekly wage regressions for white males aged 18-65.

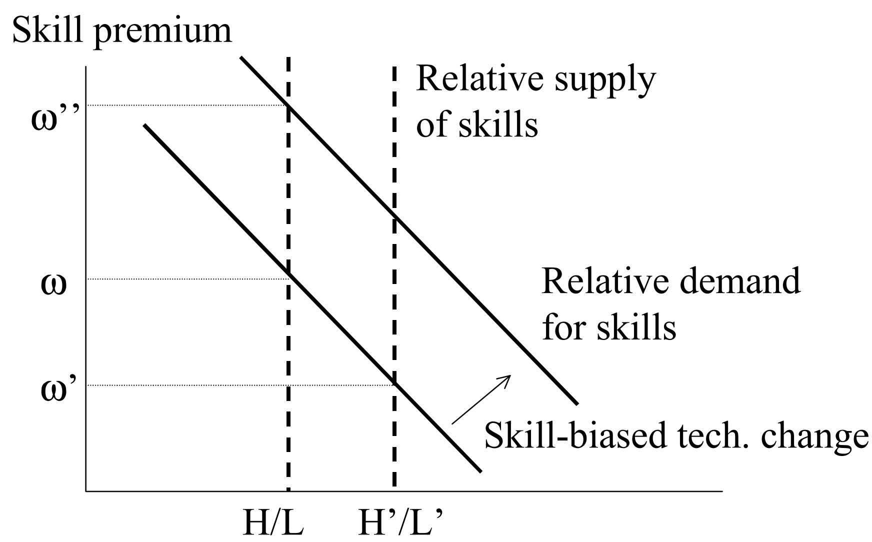

Figure 4: The relative demand for skills.

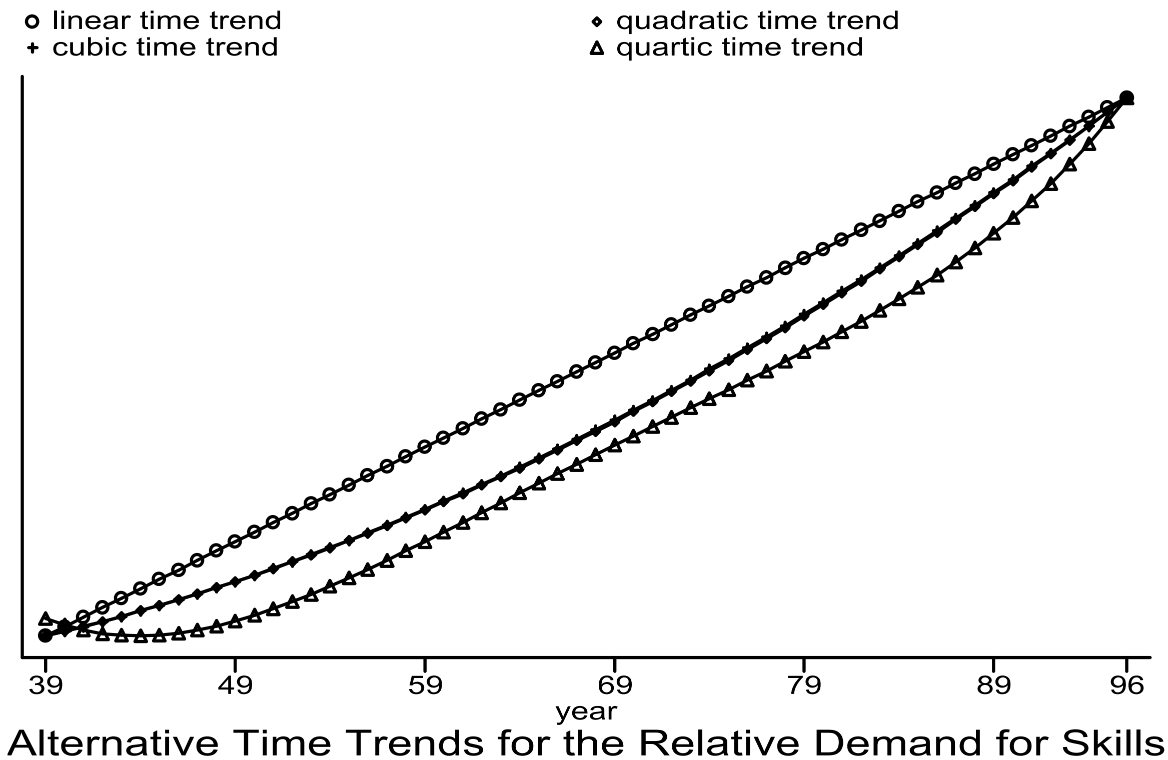

Figure 5: Estimates of time trends from regressions of $\ln \omega$ on $\ln(H/L)$, $year$, $year^2$, $year^3$ and $year^4$ between 1939 and 1996 (with observations in 1939, 1949, 1959 from the decennial censuses and observations for 1963-1996 from the March CPSs).

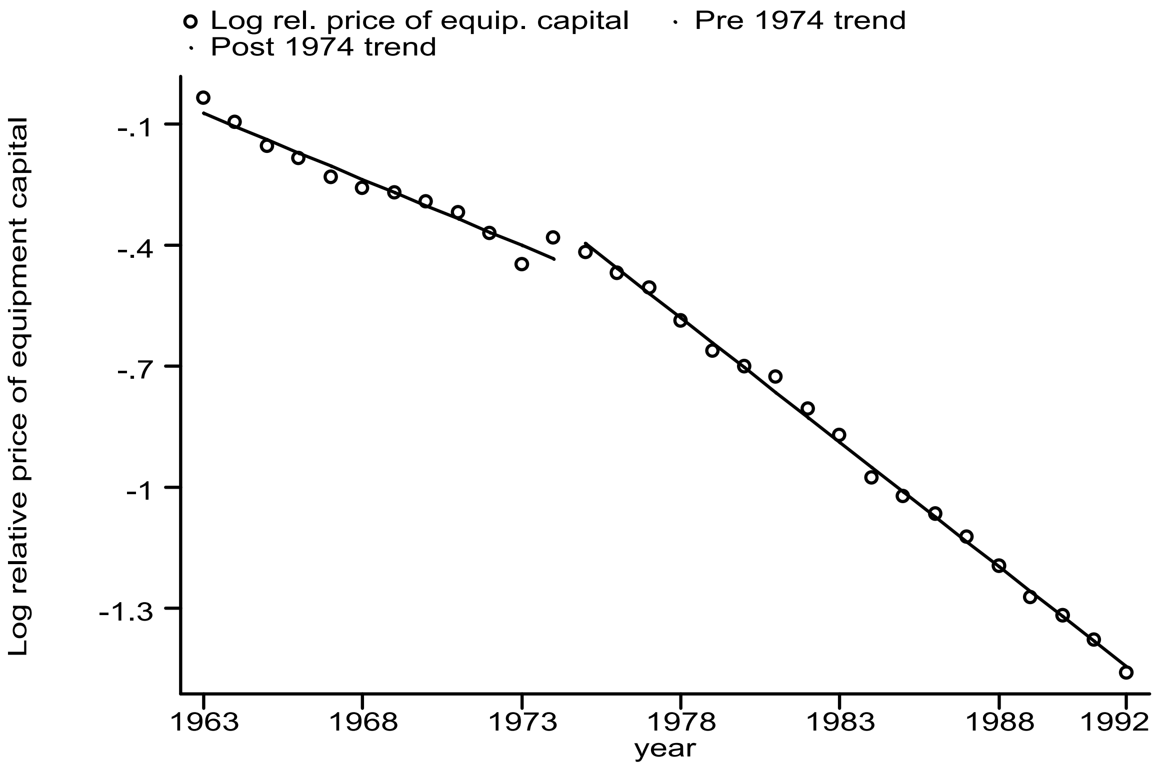

Figure 6: The behavior of the log relative price of equipment capital, 1963-1992.

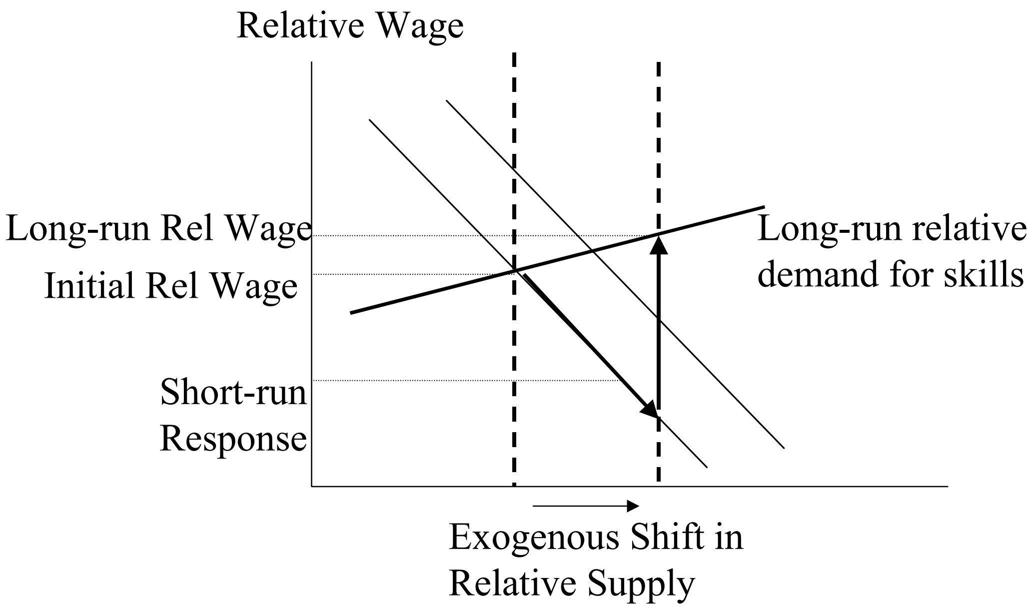

Figure 7: The dynamics of the relative wage of skilled workers in response to an increase in the supply of skills with endogenous skill-biased technical change.

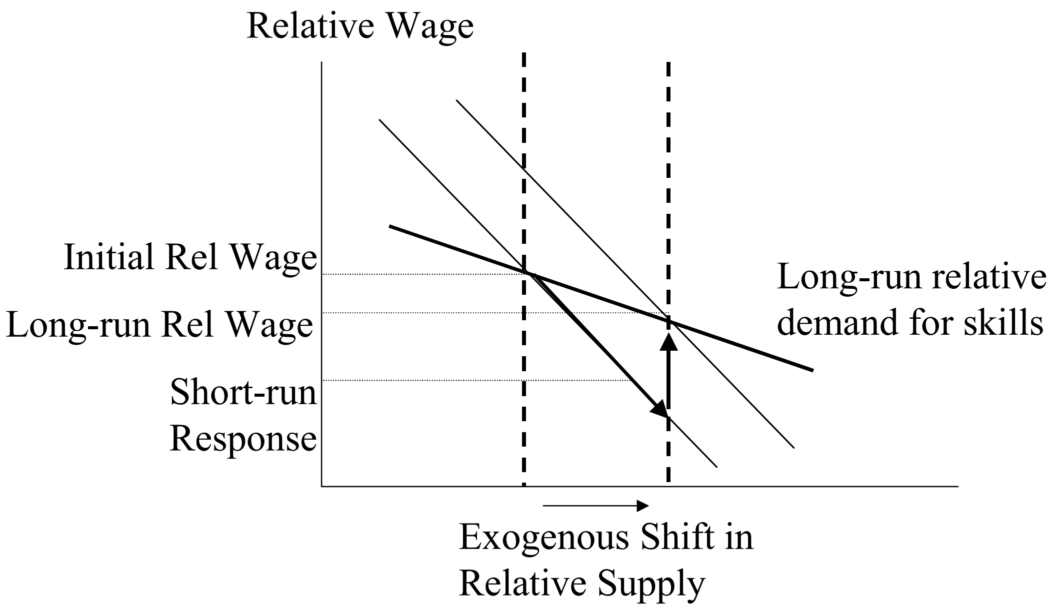

Figure 8: The dynamics of the relative wage of skilled workers in response to an increase in the supply of skills with limited endogenous skill-biased technical change.

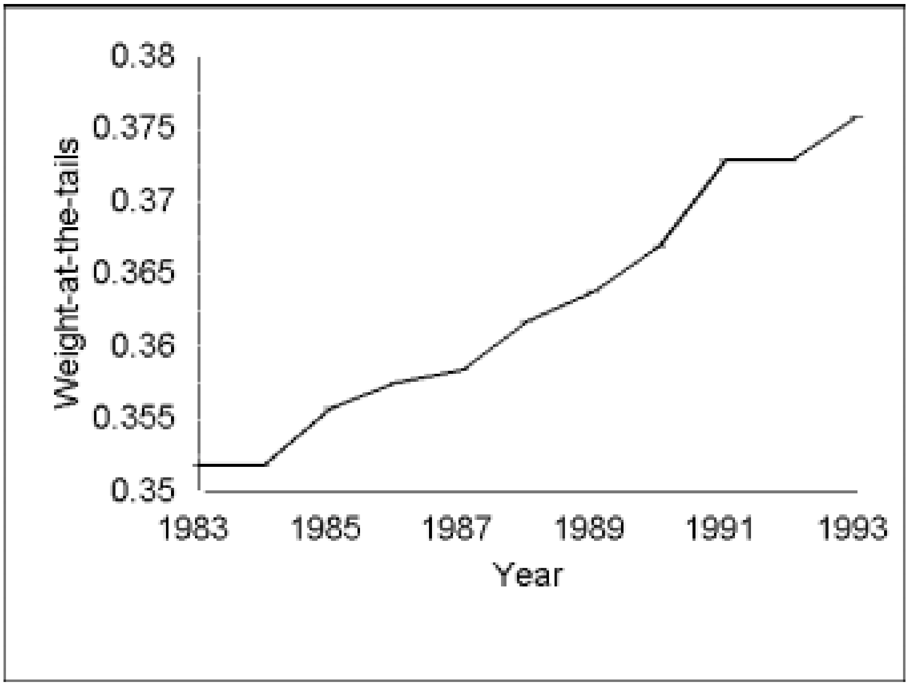

Figure 9: The evolution of the percentage of employment in the top and bottom 25 percentile industry-occupation cells (weight-at-the-tails of the job quality distribution).