14.461 Advanced Macroeconomics I: Part 3: Technology and the Structure Wages

Daron Acemoglu

November 2005

In this part I discuss the changes that have taken place in the U.S. wage structure, and also briefly discuss cross-country trends. As is well known, wage and income inequality have increased considerably in the U.S. over the past 25 years. This makes an analysis of changes in the wage structure interesting in its own right. Moreover, changes in the wage structure also imply changing labor market prices of different types of skills. Therefore, studying changes in the wage structure will be informative about the changes in the demand for different types of skills and technological developments. Finally, changes in the wage structure will also lead to different incentives for human capital investments, which we might want to understand.

1 Changes in the U.S. Wage Structure

1.1 Some basic facts

Briefly, the following are some of the major changes in the U.S. wage structure.

1. 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. This conclusion is independent of how returns to education are measured. For example, the simple linear return to schooling in a typical Mincer increased sharply. It was approximately 7.5 percent in 1980, and in 1990 it stood closer to 10 percent. But in fact, the increase is more significant between high school graduates in college graduates. Between 1979 and 1987, the average weekly wages of college graduates with one to five years of experience increased by 30 percent relative to the average weekly earnings of comparable high school graduates. The increase in inequality is even more pronounced between high school graduates and those with more than college.

2. Overall wage inequality, for example as measured by the ratio of the different percentiles of the overall wage distribution (e.g. 90-10), rose sharply beginning in the 1970s.

3. The single biggest contributor to the increase in overall wage inequality is the increase in within group (residual) inequality—i.e., increases in inequality among observationally equivalent workers. The standard way to compute residual wage inequality is either to look at inequality within very narrowly defined cells (workers with the same education level, the same experience level and of the same sex and race), or to

run a standard Mincer 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. The fact that β_t is indexed by t indicates that returns to these observed characteristics are allowed to vary from year-to-year. 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}.

4. Average and median wages have stagnated and wages of low-skill workers have fallen in real terms since 1970. 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. In the meantime, the real wages of high school graduates with 1 to 5 years of experience fell by 20 percent from 1979 to 1987.

5. These changes have been more pronounced for relatively less experienced workers, and the experience premium—the earning difference between high and low experienced workers— has also changed. In particular, among college graduates, young college graduates now earn

relatively more as compared to older college graduates than before. In contrast, among high school graduates, the earnings gap between more and less experienced workers has widened substantially.

6. The wage differential between men and women has narrowed substantially.

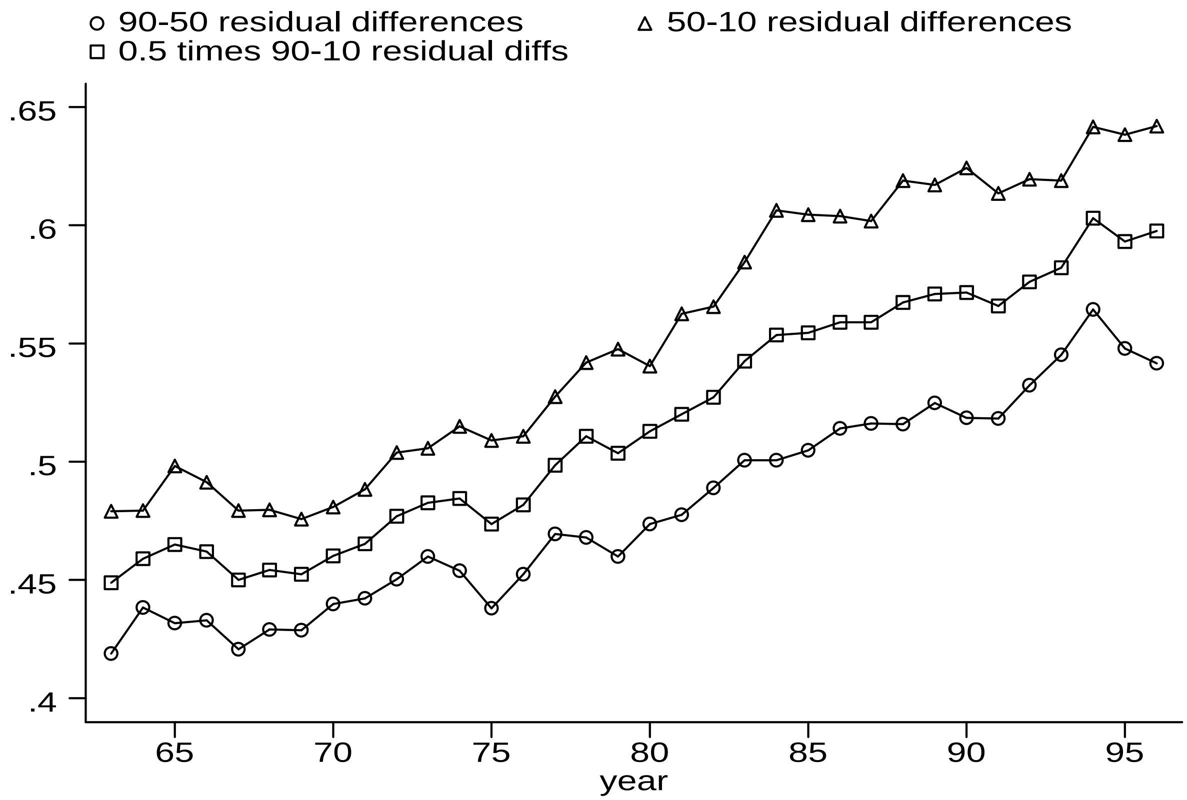
7. The wage differential between black and white workers, which had been narrowing until the mid-70s, started to expand.

8. Inequality of compensation, taking into account non-wage and fringe benefits, has expanded more than earnings inequality (Pierce, 2000).

9. Income inequality has also increased substantially over this time period, mostly reflecting the increase in wage inequality, but also the explosion in CEO pay and the high rates of return on capital and other assets which are held very unequally in the population.

10. There has been very large increase in the incomes of those at the very top of the earnings distribution (the top 1 percent or even 0.1 percent of income distribution), in part because of stock options and the very strong performance of the stock market.

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



Residual inequality measures for white males 1963-1997

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

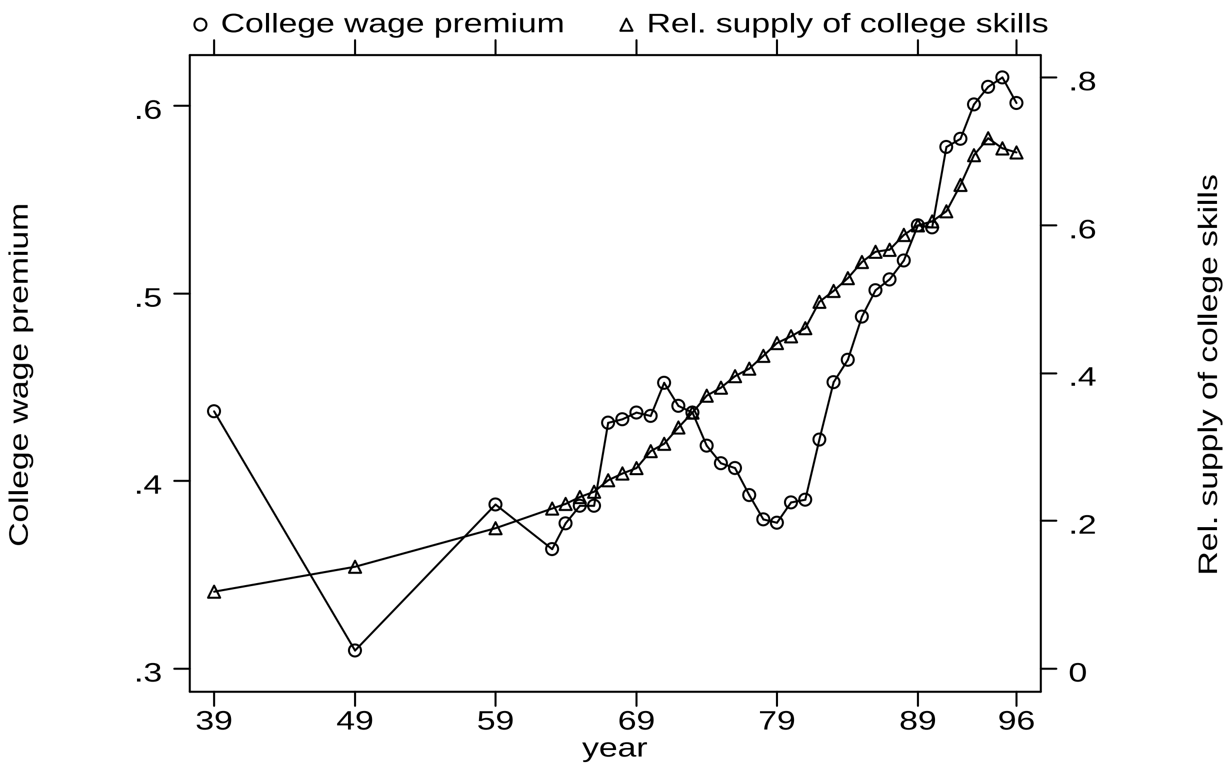
Most of these facts are not controversial. But there is some debate about two of those facts.

1. There is disagreement regarding when the increase in overall and residual inequality started. The March CPS and census data unambiguously indicate that it started in the 1970s. But May CPS data gives more ambiguous results. DiNardo, Fortin and Lemieux find that wage

inequality appears to increase starting in the 1980s in the May CPS data, but the reanalysis of these data by Katz and Autor (2000) shows consistent increases in wage inequality during the 1970s from March and May CPS data, and from census data. So I take the starting date of the increase in overall and residual inequality to be the early 1970s.

2. Some economists claim that average and median wages haven't really fallen, but this reflects mismeasurement of the CPI, which is understating wage growth. This argument is not very convincing, however. Even in the presence of such measurement problems, unless there is an "acceleration" in this bias exactly around the 1970s, there is a large gap between the rate of increase of real wages before and after the 1970s. It has to be noted, however, 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. So whether average wages have stagnated or continued to increase in line with output growth depends on how benefits are valued relative to earnings.

Finally, there is another major fact which will play an important role in the interpretation of the changes in wage inequality. There has been a remarkable increase in the supply of skills in the U.S. economy over the past sixty years, and this increase in the supply of skills accelerated starting in the early '70s. 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. Equally important, the rate of growth of the relative supply of skills significant the accelerated starting in the late 1960s, because of the Vietnam War draft laws, increase government support for education, and the high college enrollment rates of the baby boom cohorts.



Relative Supply of College Skills and College Premium

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.

1.2 Permanent versus transitory inequality

An important question for certain purposes is whether the increase in inequality is permanent or transitory (temporary). For example, if it is transitory and individuals can insure against it (for example by saving and borrowing), the increase in income and earnings inequality may not correspond to an increase in "utility" inequality.

The facts suggest that there is also an increase in the transitory component of earnings inequality, meaning that earnings fluctuations of workers around their permanent earnings levels are now greater. This fact was first noted by Gottschalk and Moffit. They document this fact by fitting an age-earnings profile for each individual, and then studying fluctuations of their incomes around this trend. The same pattern seems to arise both in the PSID and the CPS, and also can be seen in the UK.

A word of interpretation is necessary here, because the increase in the temporary component of inequality is often misinterpreted. Such an increase in temporary earnings inequality may result from two distinct causes:

1. a given worker may move up and down more frequently (greater "churning" in the labor market);

2. a worker may move up and down the same amount, but the movements may be bigger.

Mathematically, imagine the worker's "permanent" earnings level is w_0, but every period, with probability p he may get lower earnings, w_l, and with probability q, higher earnings, w_h. The two distinct causes correspond to changes in the probabilities p and q, and to changes in earnings levels w_l and w_h.

A closer look at the data indicates that all of the increase in the temporary earnings inequality is due to the second cause. For example, there is no greater movement of workers between different deciles of the wage distribution.

This pattern in fact makes a lot of sense. The leading interpretation of the fluctuations of a given worker's wages is that his skills are fluctuating over time (or the perception of the market regarding his skills are changing, for example as in Jovanovic's model). So we have $w_0 = ws_0$ where w is the market price of skills and s_0 is his regular skill level. We also have $w_h = ws_h$ and $w_l = ws_l$. If there is an increase in wage inequality corresponding to a greater skill premium, w, this will translate into a greater wage gap between different skill levels and therefore into a greater temporary component in earnings inequality. This can all happen without a change in any of the parameters p, q, s_0 or s_1 that correspond to the likelihood and magnitude of skill level fluctuations of workers.

This is important since many applied papers interpret the increase in earnings instability as reflecting greater churning. But there is basically no evidence for greater churning in the labor market. First, measures of job reallocation constructed by Davis and Haltiwanger indicate no increase in job reallocation during the past 20 or so years. Second, despite the popular perception to the contrary, there has not been a large increase in employment instability. The tenure distribution of workers today looks quite similar to what it was 20 years ago. The major exception to this seems to be middle-aged managers, who may be more likely to lose their jobs today than 20-25 years ago.

1.3 Cross-country trends

We know somewhat less about changes in wage inequality in most other countries than in the U.S. But the following are now relatively well-established:

1. Wage inequality, both returns to schooling and residual inequality, increased significantly in the UK.

2. Wage inequality also increased somewhat in other Anglo-Saxon economies, perhaps quite markedly in New Zealand.

3. In much of continental Europe, there has been little increase in either returns to schooling or overall inequality.

We return to a discussion of the possible explanations for these patterns below.

2 Interpretation

I will now focus on the two major facts related to the increase in earnings inequality: the increase in the returns to schooling and increased overall and residual inequality. In this section, I will argue two things:

1. Both of these facts may be reflecting an increase in skill premia—an increase in the market price of skills.

2. They are unlikely to reflect composition effects, cohort effects or factors related to signaling/selection.

2.1 Increase in inequality and unobserved skills

The idea that the increase in residual inequality may reflect increased skill premia is related to the idea that a large component of earnings dispersion not explained by schooling is nonetheless related to skill differences. If I am less skilled than another worker with the same education, for example, because he can perform certain tasks that I cannot, the earnings inequality between us will widen when returns to skills increase, e.g., because the tasks that he can perform become more valuable. In the data this will show up as increased residual inequality. This argument is put forth forcefully by Juhn, Murphy and Pierce (1993).

More formally, 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). Denote the unobserved skill of individual 1 by a_1 and that of individual 2 by $a_2 > a_1$, and assume that wages are given by

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

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, the variance of log wages (or of residual wages) among these two individuals is

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

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. In other words, if we ignore composition effects, which here correspondent to changes in $a_2 - a_1$, this increase must be due to a rise in the price of (and demand for) unobserved skills.

2.2 Composition effects

Could composition effects explain a significant component of the increase in residual and overall inequality, and also the bulk of the changes in the returns to education?

First, to see the reasoning for why composition effects may be important in the increase in inequality, note that the increase in the returns to education may be simply the result of an increase in the average ability of workers with high education to that of workers with low education over time. This will immediately lead to an increase in the return to education. This change in relative abilities of the two groups may result from selection (for example, today different types of workers than before may be obtaining college degrees). There is no presumption that any signaling is going on here. The employers may well be observing these skill dimensions that we, as the economists, do not observe. Alternatively, these changes in relative abilities of the two groups may result from changes in signaling behavior. College education may have become a much more important signal today, because employers may expect only the very low ability workers not to obtain a college degree. Both of these explanations could potentially account for the increase in the returns to education and residual and overall inequality.

However, the evidence suggests that the increase in the returns to education and residual inequality are not simply due to composition effects. Before discussing this evidence, note first that composition effects cannot by themselves explain the recent changes in inequality: composition effects suggest that inequality among educated and uneducated workers should move in opposite directions (see below). This suggests that changes in the true returns to skills must 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 (2) above with two education levels, high

$h = 1$ and low $h = 0$, and suppose wages are given by

$$\ln w_{it} = a_i + \gamma_t h_i + \varepsilon_{it} \tag{3}$$

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—as:

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

where $A_{1t} \equiv E\left(a_i \mid h_i = 1\right)$ and $A_{0t} \equiv E\left(a_i \mid h_i = 0\right)$. 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 relative abilities of 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, the case is not theoretically compelling: the opposite can happen as easily. 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.

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.

2.3 Evidence on composition effects

Empirically, the importance of composition effects can be uncovered by looking at inequality changes by cohort (e.g. Blackburn, Bloom and Freeman, or Juhn, Murphy and Pierce). To see this, rewrite equation (3) as

$$\ln w_{ict} = a_{ic} + \gamma_t h_{ic} + \varepsilon_{cit} \tag{4}$$

where c denotes a cohort—i.e., a group of individuals who are born in the same year, or a group of individuals who have come to the market in the same year. I have imposed an important assumption in writing equation (4): 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

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

where $A_{1ct} \equiv E\left(a_{ic} \mid h_i = 1\right)$ 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

$$\Delta \ln \omega_{c,t'-t} \equiv \ln \omega_{ct'} - \ln \omega_{ct} = \gamma_{t'} - \gamma_t, \tag{5}$$

i.e., changes in the returns to education within a cohort will reveal the true change in the returns to education.

The assumption that returns to skills are constant over the lifetime of an individual may be too restrictive, however. As we saw above, there are 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

$$\Delta^2 \ln \omega \equiv \Delta \ln \omega_{c,t'-t} - \Delta \ln \omega_{c'',t-t''} = \gamma_{t'} - \gamma_{t''}. \tag{6}$$

The evidence using this approach indicates that there are large positive changes in the returns to a college degree or this time period.

Juhn, Murphy and Pierce (1993) apply similar methodology to the increase in overall and residual inequality. They also find that these changes cannot be explained by composition effects either. 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.

(Chapter head:)The Basic Theory of Skill Premia

The simplest framework for thinking about skill premia (returns to schooling and returns to other skills) starts with a supply-demand framework. The demand for skills is often thought to be generated by the technology possibilities frontier of the economy, but is also affected by international trade, and by the organization of production.

3 The Constant Elasticity of Substitution Framework

3.1 The aggregate production function

Let me start with the simplest framework where 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, let us think of the unskilled workers as those with a high school diploma, and the skilled workers as those with a college degree, so the terms "skill" and education will be used interchangeably. In practice, however, education and skills are 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 the labor market is competitive.

The production function for the aggregate economy takes the form the constant elasticity of substitution (CES) form,

$$Y(t) = \left[(A_l(t) L(t))^{\rho} + (A_h(t) H(t))^{\rho} \right]^{1/\rho}, \tag{7}$$

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 $\sigma \equiv 1/(1-\rho)$.

Skilled and unskilled workers are 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:

1. $\sigma \to 0$ (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;

2. $\sigma \to \infty$ when skilled and unskilled workers are perfect substitutes

3. $\sigma \to 1$, 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. Depending on the value of the elasticity of substitution, an increase in A_h can act either to complement or to "replace" skilled workers.

As a little digression, at this point we can note that a more general formulation would be

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

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. For most of the analysis here, there is little to be gained from this more general production function.

The production function (7) 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 (7) 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.

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

$$w_L = \frac{\partial Y}{\partial L} = A_l^\rho \left[A_l^\rho + A_h^\rho (H/L)^\rho\right]^{(1-\rho)/\rho}. \tag{8}$$

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

$$\omega = \frac{w_H}{w_L} = \left(\frac{A_h}{A_l}\right)^\rho \left(\frac{H}{L}\right)^{-(1-\rho)} = \left(\frac{A_h}{A_l}\right)^{(\sigma-1)/\sigma} \left(\frac{H}{L}\right)^{-1/\sigma}. \tag{9}$$

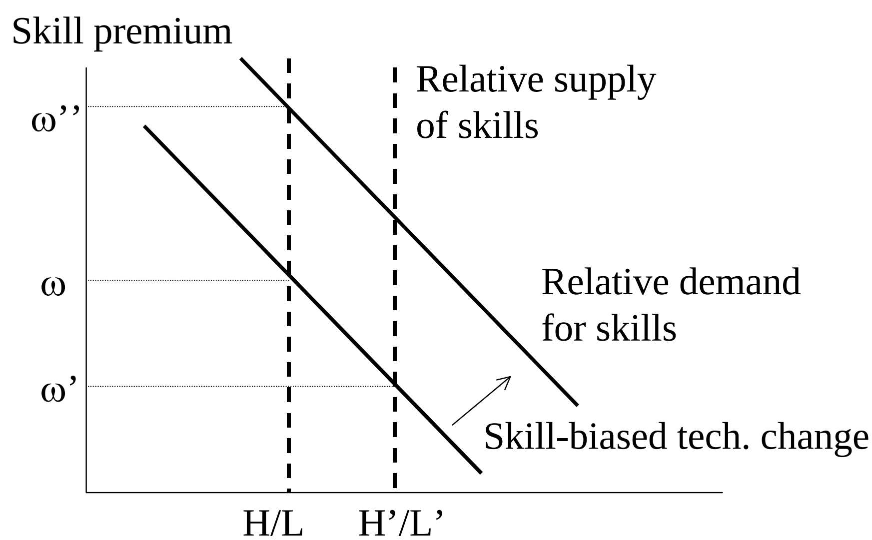
Equation (9) can be rewritten in a more convenient form by taking logs,

$$\ln \omega = \frac{\sigma - 1}{\sigma} \ln \left(\frac{A_h}{A_l}\right) - \frac{1}{\sigma} \ln \left(\frac{H}{L}\right). \tag{10}$$

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

$$\frac{\partial \ln \omega}{\partial \ln H/L} = -\frac{1}{\sigma} < 0. \tag{11}$$

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.



The relative demand for skills.

1. if skilled and unskilled workers are producing the same good, but performing different tasks, an increase in the number of skilled workers

will necessitate a substitution of skilled workers for tasks previously performed by the unskilled.

2. if skilled and unskilled workers are producing different goods (because they have different comparative advantages making them useful in different sectors), the greater number of skilled workers will lead to a substitution of the consumption of the unskill-intensive good by the skill-intensive good. In both cases, this substitution hurts the relative earnings of skilled workers.

3.2 Relative supply of skills, technology, and the skill premium

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 (11) shows, the elasticity of substitution, σ, regulates the behavior of the skill premium in response to 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. 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 $\sigma = 2$.

How does the skill premium responds to technology? Differentiation of (10) 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.

Diagrammatically, this can be seen 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 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 value over marginal product 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

$$w = \frac{Lw_L + Hw_H}{L + H} = \frac{\left[(A_l L)^\rho + (A_h H)^\rho\right]^{1/\rho}}{1 + H/L}, \tag{12}$$

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.

3.3 Summary

The results I have outlined 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, of the form

$$F\left(A_l L, A_h H, K\right)$$

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 above, the wages of low-skill workers fell over the past 30 years.

4 Why Technical Change Must Have Been Skill Biased?

The most central result for our purposes is that as H/L increases, the skill premium, ω, should fall. Diagrammatically, 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 $\frac{\sigma-1}{\sigma}\ln(A_h/A_l)$.

As discussed above, 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* to prevent the relative wages of skilled workers from declining. The cause for this steady increase in the demand for skills highlighted by this simple framework is "skill-biased technical change", broadly construed.

It is important to emphasize that skill-biased technical change here does not necessarily mean introduction of new machines that increase aggregate productivity and the relative productivity of skilled workers. Changes in organizational forms, introduction of new goods and new brands, or changes in the competitive structure of the economy may all translate into "skill-biased developments" as far as this framework is concerned. 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 $(A_h/A_l)^{(\sigma-1)/\sigma}$.

If we assume a specific value for σ, we can translate these numbers into changes in A_h/A_l to get a sense of the magnitude of the changes. In particular, notice that the relative wage bill of skilled workers is given by

$$S_H = \frac{w_H H}{w_L L} = \left(\frac{A_h}{A_l}\right)^{(\sigma-1)/\sigma} \left(\frac{H}{L}\right)^{(\sigma-1)/\sigma}. \tag{13}$$

Hence, we have

$$\frac{A_h}{A_l} = \frac{S_H^{\sigma/(\sigma-1)}}{H/L}. \tag{14}$$

We can easily 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). 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, these 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 (10) 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.

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 therefore suggests that the past sixty years must have been characterized by skill-biased technical change.

Note however that the presence of steady skill-biased technical change does not offer an explanation for the rise in inequality over the past 25 years, since we are inferring that technical change has been skill biased for much longer than these decades, and inequality was stable or even declining during the decades before the 1970s. Moreover, skill-biased technical change by itself is not enough for inequality and skill premia to increase. It will only need to increase in inequality when it outpaces the increase in the relative supply of skills. We will discuss this topic in more detail below.

(Chapter head:)Non-Technological Explanations For the Changes in the Wage Structure

Armed with a simple framework for analyzing returns to schooling and skill premia, we can now discuss what the potential causes of the changes in the way structures could be. In this section, I start with three non-technological explanations. In the next section, I will discuss theories where the increase in wage and earnings inequality may reflect changes in technologies.

By non-technological explanations I do not mean explanations in which technology plays no role, but simply that there hasn't been anything unusual in the technology front. Instead some other changes are responsible for the transformation of the wage structure. The three explanations I will discuss are:

1. The steady-demand hypothesis. According to this view, there has been no major change in the structure of technology, and therefore in the

structure of demand for skills. Changes in the returns to schooling and skill premium can be explained by the differential rates of growth in the supply of skills.

2. The trade hypothesis. According to this view, the increase in international trade, especially trade with less-developed countries, is responsible for the (unusual) increase in the demand for skills over the past twenty-five years.

3. The labor-market institutions hypothesis. This view assigns changes in the wage structure to the decline of unions, the erosion of the real value of the minimum wage and more generally, to changes in labor market regulations.

Throughout, it is useful to bear in mind that the leading alternative to the non-technology models is a view which sees an acceleration in the skill bias of technology—in other words, some "unusual" technological developments affecting the demand for skills. So I will be sometimes explicitly or implicitly comparing these three non-technological hypotheses to the technology view

5 The Steady-Demand Hypothesis

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

$$\ln\left(\frac{A_h(t)}{A_l(t)}\right) = \gamma_0 + \gamma_1 t, \tag{15}$$

where t is calendar time. Substituting this equation into (10), we obtain

$$\ln \omega = \frac{\sigma - 1}{\sigma} \gamma_0 + \frac{\sigma - 1}{\sigma} \gamma_1 t - \frac{1}{\sigma} \ln \left(\frac{H}{L} \right). \tag{16}$$

According to equation (16), the demand for skills increases 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. 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 (16) above using aggregate data between 1963-1987. 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. In fact, Katz and Murphy show that the predicted 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.

1. The regression uses only 25 aggregate observations, and there is significant serial correlation in the college premium. 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. 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.

2. 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 the numbers reported 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.

So it is important to undertake a detailed analysis of whether the steady-demand hypothesis could explain the general patterns.

6 Evidence On Steady-Demand Vs. 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. Equally, however, it would be wrong to interpret the findings of DiNardo and Pischke (1997) and Enhorf and Kramartz (1998) as evidence against an 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.

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 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, or those with greater R&D expenditure and faster productivity increases, 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 (14). 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 above, the only way to make sense of post-war trends is to incorporate skill-biased technical change over the whole period. 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 is hard to answer 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. In fact, when Autor Katz and Krueger (1998) run the above 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.

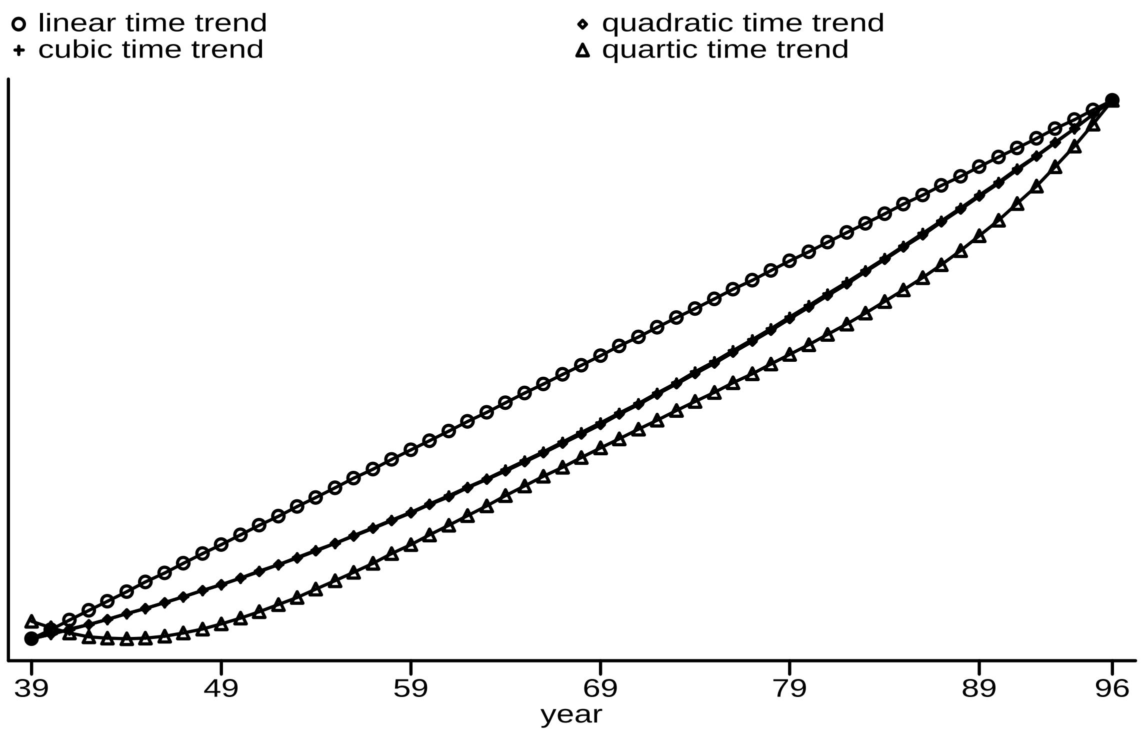
The third, and probably most powerful, piece of evidence also comes from Autor, Katz and Krueger (1998). They document that the supply of skills grew faster between 1970 and 1995 than between 1940 and 1970—by 3.06

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. 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. 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. The next figure shows these higher order trends, which indicate an acceleration in skill bias starting in the 1970s.



Alternative Time Trends for the Relative Demand for Skills

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).

A final piece of evidence comes from the behavior of overall and residual inequality over the past several decades. As argued above, this increase in inequality weighs in favor of a marked change in labor market prices and demand for skills.

Overall, therefore, 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.

It is useful to bear in mind, however, that the unusual increase in the demand for skills might be non-technological. It might reflect effect of increased international trade with skill-scarce countries, or it may reflect the collapse of some labor market institutions.

7 Trade and Inequality

7.1 Trade and wage inequality: theory

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 above. 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

$$p^{US} = \frac{p_h}{p_l} = \left[\frac{A_h H}{A_l L}\right]^{\rho - 1}. \tag{17}$$

The skill premium is then simply equal to the ratio of the marginal value products of the two types of workers, that is,

$$\omega^{US} = p^{US}\frac{A_h}{A_l} \tag{18}$$

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-intensive and labor-intensive goods are $A_h\left(H + \widehat{H}\right)$ and $A_l\left(L + \widehat{L}\right)$, the relative price of the skill intensive good will be

$$p^W = \left[\frac{A_h\left(H + \widehat{H}\right)}{A_l\left(L + \widehat{L}\right)}\right]^{\rho - 1} > p^{US}. \tag{19}$$

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, pushing 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

$$\omega^W = p^W \frac{A_h}{A_l} > \omega^{US} \tag{20}$$

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 that have started trading more with the U.S. or other OECD economies.

7.2 Evidence

Although this analysis shows that increased international trade could be responsible for the rise in skill premia and inequality in the U.S., most economists discount the role of trade for a variety of reasons.

First, as the comparison of equations (18) and (20) shows, the effect of international trade works through a *unique intervening mechanism*: free trade with the LDCs increases the relative price of skill-intensive goods, p, and affects the skill premium via this channel. 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 found that during the 1980s the relative price of skill-intensive goods actually fell. Sachs and Shatz found no major change or a slight decline, while a more recent paper by Krueger found an increase in the relative price of skill-intensive goods, but only for the 1989-95 period.

Second, as pointed out above, a variety of evidence suggests that all sectors, even those producing less skill-intensive goods, increased their demands for more educated workers. This pattern is consistent with the importance of skill-biased technical change, but not with an increase in the demand for skills driven mainly by increased international trade.

Third, as noted above, a direct implication of the trade view is that, 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. 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 illustrates this point by undertaking a calibration of a simple North-South model. Katz and Murphy, Berman, Bound and Griliches and Borjas, Freeman and Katz emphasize the same point by showing that the content of unskilled labor embedded in U.S. imports is small relative to the changes in the supply of skills taking place during this period.

These arguments suggest that increased international trade with the LDCs is not the major cause of the changes in the wage structure by itself.

8 Labor Market Institutions and Inequality

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.

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 in overall inequality for a number of reasons:

1. 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.

2. The difference between the 90th percentile and the median mirrors the behavior of the difference between the median and the 10th percentile. (Perhaps with the exception of 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). 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.

3. Perhaps most importantly, the erosion in the real value of the minimum wage started in the 1980s, whereas, as shown above, the explosion in overall wage inequality 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. Throughout the postwar period in the U.S. economy, unions negotiated the wages for many occupations, even indirectly influenced managerial salaries. Unions also explicitly tried to compress wage differentials. This suggests that the decline of unions may be a major cause of the changes in the structure of wages.

But once again, deunionization does not appear to be the major cause of the increase in inequality.

1. Wage inequality increased in many occupations in which prices were never affected by unions (such as lawyers and doctors)—and importantly, these are occupations with highly specific skills, so there is no presumption that the increase in inequality in other occupations should directly translate to these occupations.

2. Perhaps more important, 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. 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. In Canada, while unionization rates increased from around 30 percent in the 1960s to over 36 percent in the late 1980s, wage inequality also increased.

(Chapter head:)Acceleration in Skill Bias

By the process of elimination, we have arrived at an acceleration in skill bias as the most likely cause of the increase in inequality over the past 25-30 years.

There is also a variety of direct evidence that also suggests that the technological developments of the past 30 years may have increased the demand for skills considerably. Most observers agree that many computer-based technologies require a variety of abstract and problem-solving skills, thus increasing the demand for college education and related skills. The evidence in Autor, Katz and Krueger indicates that industries that have invested in computer technology have increased their demand for skills substantially. They also show that this correlation is not driven by some omitted factors, such as R&D investment or capital intensity.

9 Exogenous Acceleration Skill Bias

The most popular view in the literature is that the past 25 years experienced an acceleration in skill bias because of exogenous technological developments, and link these technological developments to a "technological revolution," most likely associated with the microchip, the computer technology, and improvements in communications technology.

I would like to distinguish between 3 different approaches here.

9.1 Exogenous skill-technology complementarity

This is the simplest view, and claims that for some (unknown/exogenous) reason there has been a more rapid increase in A_h/A_l during this period, translating into greater skill premia. This may be linked to the introduction of computer technology, but in this view there is no explicit theory of why it would be so. In other words, new technologies just happened to be more skill-biased and increase A_h/A_l.

The advantage of this theory is its simplicity, and the disadvantage is that by a similar approach we could explain anything that happens. So for this theory to make progress, that has to be more empirical work somehow documenting that A_h/A_l has increased. Cross-industry regressions that explain the demand for skills by computer investment come close to this, but are not entirely satisfactory or easy to interpret.

An interesting attempt in this direction is made by Autor, Levy and Murnane. They argue that the recent increase in the skill bias is an outcome of improvements in computer technology resulting from the fact that computers substitute for routine tasks, while at the same time increasing the demand for problem-solving skills. Using data from the Dictionary of Occupational Titles and decennial censuses, they show that in industries with greater computerization, there has been a shift away from occupations specializing in routine tasks towards occupations with a heavy problem-solving component. Therefore, this approach gives some empirical content to the otherwise exogenous and hard-to-observe skill bias. It also suggests that the recent acceleration in the demand for skills may be the outcome of the nature of computers, which by construction replaced routine, easy-to-replicate tasks.

9.2 Capital-skill complementarity

An interesting fact uncovered first by Gordon is that the relative price of equipment capital has been falling steadily in the postwar period. Moreover, this rate of decline of equipment prices may have accelerated sometime during the mid to late 1970s.

Krusell, Ohanian, Rios-Rull and Violante, using this fact, argue that the demand for skills accelerated as a result of the more rapid decline in the relative price of capital equipment. They build on an idea first suggested by Griliches that capital is more complementary to skilled rather than unskilled labor. Combining this insight with the relatively rapid accumulation of equipment capital due to the decline its relative price, they offerr an explanation for why the demand for skills may have increased at a faster rate during the past 25 years than before.

More explicitly, these authors consider the following aggregate production function

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

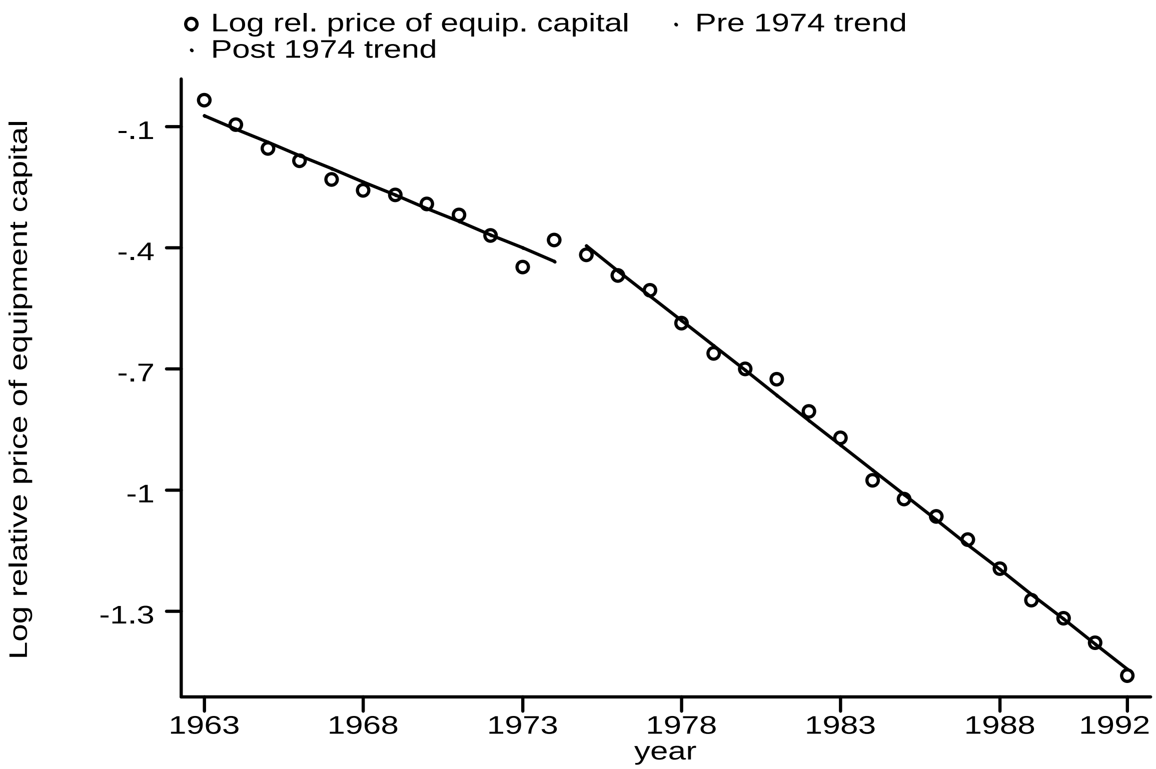
where K_s is structures capital (such as buildings), and K_e is equipment capital (such as machines). The parameter $\sigma_1 = 1/(1-\lambda)$ is the elasticity of substitution between equipment and skilled workers, and $\sigma_2 = 1/(1-\mu)$ is the elasticity of substitution between unskilled workers 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.

The skill premium in this model is

$$\omega = \frac{(1-b_2)(1-b_1)H^{\lambda-1}\left(b_2 K_e^{\lambda} + (1-b_2)H^{\lambda}\right)^{(\mu-\lambda)/\lambda}}{b_1 L^{\mu-1}}$$

Differentiation 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.



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

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 (16) 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.

But, the evidence suggests that the relative quantity of equipment capital or its relative price never does as well as a time trend. When entered together in a time-series regression, 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.

9.3 Technological revolutions and the Schultz view of human capital

Recall that according to Schultz/Nelson-Phelps view of human capital, skills and ability are more useful at times of rapid change (at times of "disequilibrium" as Schultz called it). So if indeed there has been a technological revolution, we might expect this may have increased the demand for skills.

This view has been advanced by a number of authors, including Greenwood and Yorukoglu (1997), Caselli (1999), Galor and Moav (2000). For example, Greenwood and Yorukoglu (1997, p. 87) argue:

> "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."

Let me give a brief formalization of this approach built on Galor and Moav (2000) adapted to the above framework. Suppose that in terms of the CES framework developed above

$$A_l = \phi_l(g)a \text{ and } A_h = \phi_h a \tag{21}$$

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 (21) into (9), the skill premium is

$$\omega = \frac{w_H}{w_L} = \left(\frac{A_h}{A_l}\right)^{(\sigma-1)/\sigma} \left(\frac{H}{L}\right)^{-1/\sigma} = \left(\frac{\phi_h}{\phi_l(g)}\right)^{(\sigma-1)/\sigma} \left(\frac{H}{L}\right)^{-1/\sigma}. \tag{22}$$

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

This approach therefore presumes that the recent past has been characterized by faster than usual technological progress, and explains the acceleration in skill bias by the direct effect of more rapids technical change on the demand for skills.

10 Review of Models of Endogenous Technical Change

There are two basic types of models of endogenous technical change that lead to steady growth. There are also many other possible formulations, but these models do not generally lead to steady growth, so have received less attention. Moreover, models that lead to steady growth are easier to analyze, since they have a quasi-linear structure. Here I will focus on the class of models that leads to steady growth.

The important feature of the models of endogenous technology in general for the purposes of these notes is that they endogenize the productive capacity of an economy, and provide a framework for analyzing how a variety of factors may affect productivity.

10.1 Models of Expanding Variety

Models of expanding variety were first introduced by Dixit and Stiglitz, and then used extensively in the analysis of international trade. For this reason, they are often formulated as a situation in which "product" variety expands. For the purposes of this course, it is easier to think of expanding variety of inputs. So let us consider the following model.

Imagine an infinite-horizon representative agent economy

$$\max \int_0^\infty \frac{C^{1-\theta} - 1}{1-\theta} \cdot e^{-\rho \cdot t} \cdot dt$$

The unique consumption good of the economy is produced as

$$Y = \frac{1}{1-\beta}\left[\int_0^N k(v)^{1-\beta} dv\right] \cdot L^{\beta} \tag{23}$$

where L is the aggregate labor input, N denotes the different number of varieties of capital inputs, and k is the total amount of capital (machine) of input type v.

The budget constraint of the economy is

$$C + I + X \leq Y \tag{24}$$

where I is investment and X is expenditure on R&D.

Assume that

$$\dot{N} = X \tag{25}$$

so R&D expenditure expands the potential set of capital varieties. A firm that invents a new capital variety is the sole supplier of that type of machine, and sets its price $\chi(v)$ to maximize profits. The demand for capital of type v is obtained by maximizing (23) and takes the convenient isoelastic form:

$$k(v) = \left[L^{\beta}/\chi(v)\right]^{1/\beta} \tag{26}$$

Machines depreciate fully after use and creating one unit of machine costs ψ units of output.

Consider the monopolist owning a machine ν invented at time 0. This monopolist chooses an investment plan and a sequence of capital stocks so as to maximize the present discounted value of profits, as given by

$$V(\nu) = \int_0^{\infty} e^{-rt}\left[\chi(\nu)k(\nu) - \psi k(\nu)\right] dt$$

Alternatively, this could be written as a dynamic programming equation of the form

$$rV(\nu) - \dot{V}(\nu) = \chi(\nu)k(\nu) - \psi k(\nu)$$

Since (26) defines isoelastic demands, the solution to this program involves

$$\chi(\nu) = \psi/(1-\beta)$$

that is, all monopolists charge a constant rental rate, equal to a mark-up over the marginal cost times. Without loss of generality, normalize the marginal cost of machine production to $\psi \equiv (1-\beta)$, so that

$$\chi = 1$$

Profit-maximization also implies that each monopolist rents out the same quantity of machines in every period, equal to $k(v) = L$, and makes profits

$$\pi = \beta L \tag{27}$$

Substituting (26) and the machine prices into (23), we obtain

$$Y = \frac{1}{1-\beta} NL$$

so output grows at the same rate as the number of varieties.

Therefore, to determine the rate of growth of output, we only need to determine the rate of growth of N. This can happen via three different processes.

1. R&D by the economy in question.

2. Imitation from the R&D or production techniques of other countries.

3. Exogenously.

Here I choose the first possibility, as specified by the technology possibilities of the economy as in (25)

Let us focus on the first channel for now. From the above analysis, the steady-state value of an invention (when $\dot{V} = 0$):

$$V = \frac{\pi}{r}$$

where π is the flow of net profits per period, given by (27) above.

For there not to be further incentives to undertake R&D, we need

$$\frac{\beta L}{r} = 1$$

This equation pins down the steady-state interest rate as:

$$r = \beta L$$

From consumer maximization, we also have that the rate of growth of consumption, g_c, is given by

$$g_c = \frac{1}{\theta}(r - \rho) \tag{28}$$

and in steady state, the rate of growth of the economy is the same as the rate of growth of consumption, so we have that the whole economy grows at the rate g.

Therefore, given the steady-state interest rate we can simply determined the long-run growth rate of the economy as:

$$g = \frac{1}{\theta}(\beta L - \rho)$$

Notice that there is a scale effect here: the larger is L, the greater is the growth rate. The scale effect comes from the increasing returns to scale nature of the technology of model of endogenous technical change (this is a point related to the non-rival nature of knowledge, emphasized in Romer, 1990).

Jones in a series of papers has argued that there is little evidence for such scale effects, and has suggested generalizations of these models to kill the scale effect. For the purposes here, the scale effect is not of primary importance.

10.2 Growth of Knowledge Spillovers

In the model of the previous section, growth resulted from the use of final output for R&D. This is similar, in some way, to the endogenous growth model of Rebelo, since the accumulation equation is linear in accumulable factors.

An alternative is to have "scarce factors" used in R&D. In this case, there will not be endogenous growth, unless there are knowledge spillovers from past R&D. In other words, now current researchers need to "stand on the shoulder of past giants".

A typical formulation in this case is

$$\dot{N} = N \cdot L_R \tag{29}$$

where L_R is labor allocated to R&D. The term N on the right-hand side captures spillovers from the stock of existing ideas. The greater is N, the more productive is an R&D worker.

L_R could be skilled workers as in Romer (1990), or scientists or regular workers. In the latter case, there will be competition between the production sector and the R&D sector for workers, and the marginal cost of workers and research would be given by the wage rate and production sector. In particular, the free entry condition is now

$$N \cdot V = w$$

where N is on the left-hand side because it parameterizes the productivity of an R&D worker.

For example, in the model I outlined above, the wage rate is

$$w = \frac{\beta}{1-\beta} N$$

So the steady-state free-entry condition becomes

$$N \frac{\beta L}{r} = \frac{\beta}{1-\beta} N$$

Hence the long-run equilibrium interest-rate is

$$r = (1-\beta) L$$

The rest of the analysis is unchanged.

10.3 Models of Quality Competition

In the model of expanding machine variety, different machines were complements in production. However, in practice when a better computer make comes to the market, it replaces previous models. This is captured in the models of vertical quality competition, such as the models in Aghion and Howitt, or Grossman and Helpman. The major difference from the previous setup is that the production function is now

$$Y = \frac{1}{1-\beta}\left[\int_0^1 q(v)k(v)^{1-\beta}dv\right] \cdot L^{\beta} \tag{30}$$

where $q(v)$ is the quality of machine v, and because now the number of varieties is constant, I have normalized it to 1.

To invent a new machine, firms undertake R&D on an existing machine (of type v). If a firm spends $q \cdot z$ units of the final good for R&D on a machine of quality q, then it has a probability $\mu \cdot z$ of inventing a new machine, with quality λq. The new machine will take over the market for this type of capital, but unless λ is very large, it will have to charge a limit price in order to exclude the previous leader. I assume that λ is not too large, so we will observe limited prices in equilibrium. Again, denote the marginal cost of production is $\psi \cdot q$ for a machine of quality q.

The demand for machines are now

$$k(v) = \left[q(v)/\chi(v)\right]^{1/\beta} L \tag{31}$$

I normalize $\psi = \lambda^{-1}$, so the monopolist sets the price $\chi = 1$, and sells

$k(v) = L$. This generates profits

$$\pi(v) = \frac{\lambda - 1}{\lambda} L$$

Substituting (31) into (30), we obtain total output as

$$Y = \frac{1}{1-\beta} QL$$

where

$$Q = \int_0^1 q(v) dv$$

is the average total quality of machines.

The value of being the inventor is different now, because this position will not last forever. More formally, the standard dynamic programming equation is

$$rV - \dot{V} = \pi - xV$$

where x is the rate at which new innovations occur.

Note that there is a relationship between the innovation rate, x, and the growth rate, g, given by:

$$g = (\lambda - 1) x$$

Free entry into R&D at the cost μ implies that

$$V = \mu$$

Otherwise, there will be entry into or exit from research.

In steady state, $\dot{V} = 0$. So

$$V = \frac{\pi}{r + g/(\lambda - 1)} = \frac{(\lambda - 1)^2 L}{\lambda[(\lambda - 1)r + g]} = \mu$$

where the last equality follows from free entry. To see this, notice that one more unit of the final good provides a flow rate μ of obtaining V. Moreover, equation (28) still applies, so combining those, we have that in steady state, $r = \theta g + \rho$, so

$$\frac{(\lambda-1)^2 L}{\lambda\left[(\lambda-1)\left(\theta g+\rho\right)+g\right]} = \mu$$

therefore

$$g = \frac{1}{\lambda((\lambda-1)\theta+1)}\left((\lambda-1)^2\mu^{-1}L - (\lambda-1)\lambda\rho\right)$$

11 Directed Technical Change

The framework analyzed so far assumed technical change to be neutral towards different factors, and in fact, in most applications, we limited ourselves to the Cobb-Douglas production function.

Technical change is often not neutral towards different factors of production, and the elasticity of substitution between different factors is often found not to be equal to 1.

So it is important to consider the implications of more general production functions, and think of endogenizing technology and technological differences within this more general framework. There are, however, reasons for economists' focus on Cobb-Douglas production function. The most important one is that a general production function, associated with arbitrary technological progress, does not generate balanced growth. Instead, with a non-Cobb-Douglas production function, balanced growth requires all technical change to be labor-augmenting. Therefore, once we abandon the Cobb-Douglas production function, we need to develop a theory of why technical change is purely labor-augmenting, and a more generally think about various biases in the nature of technical change.

Do we have reason to think that biased technical change is important? The answer appears to be yes—there are many examples of systematic biases in technical change. For example, the consensus among labor and macroeconomists is that technical change throughout the 20th century has been skill-biased. There is also a possible acceleration in skill-biased technical change during the past 25 years. In contrast, evidence suggests that technical change change during the 19th century may have been, at least in part, skill-replacing.

Question: What explains these various biases and the direction of technical change?

Let us consider a model in which profit incentives determine what type of technologies are developed. When developing technologies complementing a particular factor (say skilled workers) is more profitable, more of these technologies will be developed. Whether the development of these technologies makes aggregate technology more skill-biased or not will depend on the elasticity of substitution between this factor and the rest.

What determines the relative profitability of developing different technologies?

1. The price effect: there will be stronger incentives to develop technologies when the goods produced by these technologies command higher prices.

2. The market size effect: it is more profitable to develop technologies that have a larger market.

   Schmookler (1966): "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."

11.1 Definitions

First consider what factor-augmenting and factor-biased technical change correspond to. For this purpose, take the standard the constant elasticity of substitution (CES) production function

$$y = \left[\gamma \left(A_L L\right)^{\frac{\sigma-1}{\sigma}} + (1-\gamma)\left(A_Z Z\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}},$$

where L is labor, and Z denotes another factor of production, which could be capital or skilled labor.

Here $\sigma \in (0,\infty)$ is the elasticity of substitution between the two factors.

A_L is labor-augmenting (labor-complementary) and A_Z is Z-complementary. The relative marginal product of the two factors:

$$\frac{MP_Z}{MP_L} = \frac{1-\gamma}{\gamma}\left(\frac{A_Z}{A_L}\right)^{\frac{\sigma-1}{\sigma}}\left(\frac{Z}{L}\right)^{-\frac{1}{\sigma}}. \tag{32}$$

This implies that when $\sigma > 1$, i.e., when the two factors are gross substitutes, A_L is labor-biased and A_Z is Z-biased. In contrast, when $\sigma < 1$, i.e., when the two factors are gross complements, A_Z is labor-biased and A_L is Z-biased.

11.2 Basic Model

Now we are in a position to consider a simple model of directed technical change. Assume that preferences are again given by the CRRA function

$$\int_0^\infty \frac{C^{1-\theta}-1}{1-\theta} e^{-\rho t} dt, \tag{33}$$

The budget constraint:

$$C + I + R \leq Y \equiv \left[\gamma Y_L^{\frac{\varepsilon-1}{\varepsilon}} + (1-\gamma) Y_Z^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}} \tag{34}$$

In words, the output aggregate is produced from two other (intermediate) goods, Y_L and Y_Z, with elasticity of substitution ε. Here Y can either be interpreted as the final good aggregated from the two intermediates, Y_L and Y_Z, or Y could be an index of utility defined over the two final goods, Y_L and Y_Z.

Intermediate good production functions are:

$$Y_L = \frac{1}{1-\beta} \left(\int_0^{N_L} x_L(j)^{1-\beta} dj \right) L^\beta, \tag{35}$$

and

$$Y_Z = \frac{1}{1-\beta} \left(\int_0^{N_Z} x_Z(j)^{1-\beta} dj \right) Z^\beta, \tag{36}$$

Assume that machines to both sectors are supplied by "technology monopolists". This is a straightforward generalization of the endogenous technical change model of product variety discussed above.

Each monopolist sets a rental price $\chi_L(j)$ or $\chi_Z(j)$ for the machine it supplies to the market.

The marginal cost of production is the same for all machines and equal to $\psi \equiv 1-\beta$ in terms of the final good.

Price taking implies

$$\max_{L,\{x_L(j)\}} p_L Y_L - w_L L - \int_0^{N_L} \chi_L(j)\, x_L(j)\, dj, \tag{37}$$

This gives machine demands as

$$x_L(j) = \left[\frac{p_L}{\chi_L(j)}\right]^{1/\beta} L. \tag{38}$$

Similarly

$$x_Z(j) = \left[\frac{p_Z}{\chi_Z(j)}\right]^{1/\beta} Z, \tag{39}$$

Since the demand curve for machines facing the monopolist, (38), is isoelastic, the profit-maximizing price will be a constant markup over marginal cost. In particular, all machine prices will be given by

$$\chi_L(j) = \chi_Z(j) = 1.$$

Profits of technology monopolists are obtained as

$$\pi_L = \beta p_L^{1/\beta} L \text{ and } \pi_Z = \beta p_Z^{1/\beta} Z. \tag{40}$$

Let V_Z and V_L be the net present discounted values of new innovations. Then in steady state:

$$V_L = \frac{\beta p_L^{1/\beta} L}{r} \text{ and } V_Z = \frac{\beta p_Z^{1/\beta} Z}{r}. \tag{41}$$

The greater is V_Z relative to V_L, the greater are the incentives to develop Z-complementary machines, N_Z, rather than N_L.

This highlights the two effects on the direction of technical change that I mentioned above:

1. The price effect: a greater incentive to invent technologies producing more expensive goods.

2. The market size effect: a larger market for the technology leads to more innovation. The market size effect encourages innovation for the more abundant factor.

Substituting for relative prices, relative profitability is

$$\frac{V_Z}{V_L} = \left(\frac{1-\gamma}{\gamma}\right)^{\frac{\varepsilon}{\sigma}} \left(\frac{N_Z}{N_L}\right)^{-\frac{1}{\sigma}} \left(\frac{Z}{L}\right)^{\frac{\sigma-1}{\sigma}}. \tag{42}$$

where

$$\sigma \equiv \varepsilon - (\varepsilon - 1)(1-\beta).$$

is the (derived) elasticity of substitution between the two factors. An increase in the relative factor supply, Z/L, will increase V_Z/V_L as long as $\sigma > 1$ and it will reduce it if $\sigma < 1$.

Therefore, the elasticity of substitution regulates whether the price effect dominates the market size effect.

Note also that we have

$$\sigma \geq 1 \iff \varepsilon \geq 1$$

So the two factors will be gross substitutes when the two goods in utility function (or the two intermediates in the production of the final good) are gross substitutes.

Also note that relative factor prices are

$$\frac{w_Z}{w_L} = p^{1/\beta}\frac{N_Z}{N_L} = \left(\frac{1-\gamma}{\gamma}\right)^{\frac{\varepsilon}{\sigma}} \left(\frac{N_Z}{N_L}\right)^{-\frac{\sigma-1}{\sigma}} \left(\frac{Z}{L}\right)^{-\frac{1}{\sigma}}. \tag{43}$$

First, the relative factor reward, w_Z/w_L, is decreasing in the relative factor supply, Z/L.

Second, the same combination of parameters, $\frac{\sigma-1}{\sigma}$, which determines whether innovation for more abundant factors is more profitable also determines whether a greater N_Z/N_L —i.e., a greater relative physical productivity of factor Z— increases w_Z/w_L

When $\sigma > 1$, greater N_Z/N_L increases w_Z/w_L, but when $\sigma < 1$, it has the opposite effect.

Implication: greater Z/L always causes Z-biased technical change.

We have so far characterized the demand for new technologies. Next we have to determine the "supply" all new technologies, which will be, in part, regulated by the technological possibilities for generating new machine varieties. Suppose

$$\dot{N}_L = \eta_L X_L \text{ and } \dot{N}_Z = \eta_Z X_Z, \tag{44}$$

where X denotes R&D expenditure.

This gives the following "technology market clearing" condition:

$$\eta_L \pi_L = \eta_Z \pi_{Z.} \tag{45}$$

Then, relative physical productivities can be solved for

$$\frac{N_Z}{N_L} = \eta^{\sigma} \left(\frac{1-\gamma}{\gamma} \right)^{\varepsilon} \left(\frac{Z}{L} \right)^{\sigma-1}. \tag{46}$$

Relative factor rewards are

$$\frac{w_Z}{w_L} = \eta^{\sigma-1} \left(\frac{1-\gamma}{\gamma} \right)^{\varepsilon} \left(\frac{Z}{L} \right)^{\sigma-2}. \tag{47}$$

Comparing this equation to the relative demand for a given technology, we see that the response of relative factor rewards to changes in relative supply is always more elastic in (47) than in (43).

This is simply an application of the LeChatelier principle, which states that demand curves become more elastic when other factors adjust—that is, the relative demand curves become flatter when "technology" adjusts.

The more important and surprising result here is that if σ is sufficiently large, in particular if $\sigma > 2$, the relationship between relative factor supplies and relative factor rewards can be upward sloping.

11.3 Implications

If $\sigma > 2$, then the long-run relationship between the relative supply of skills and the skill premium is positive.

To see why this is interesting and potentially important, recall three salient patterns about demand for skills:

1. Secular skill-biased technical change increasing the demand for skills throughout 20th century.

2. Possible acceleration in skill-biased technical change over the past 25 years.

3. Many skill-replacing technologies during the 19th century.

With an upward sloping relative demand curve, or simply with the degree of skilled bias endogenized, we have a natural explanation for all of these patterns.

1. The increase in the number of skilled workers that has taken place throughout 20th century is predicted to cause steady skill-biased technical change.

2. Acceleration in the increase in the number of skilled workers over the past 25 years is predicted to induce an acceleration in skill-biased technical change.

3. Large increase in the number of unskilled workers available to be employed in the factories during the 19th century could be expected to induce skill-replacing/labor-biased technical change.

In addition, this framework with endogenous technology also gives a nice interpretation for the dynamics of the college premium during the 1970s and 1980s. It is reasonable to presume 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 in the figure. 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 approach can therefore explain 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.

If on the other hand we have $\sigma < 2$, the long-run relative demand curve will be downward sloping, though again it will be shallower than the short-run relative demand curve. Then following the increase in the relative supply of skills there will be an initial decline in the skill premium (college premium), and as technology starts adjusting the skill premium will increase. But it will end up below its initial level. To explain the larger increase in the 1980s, in this case we need some exogenous skill-biased technical change. The next

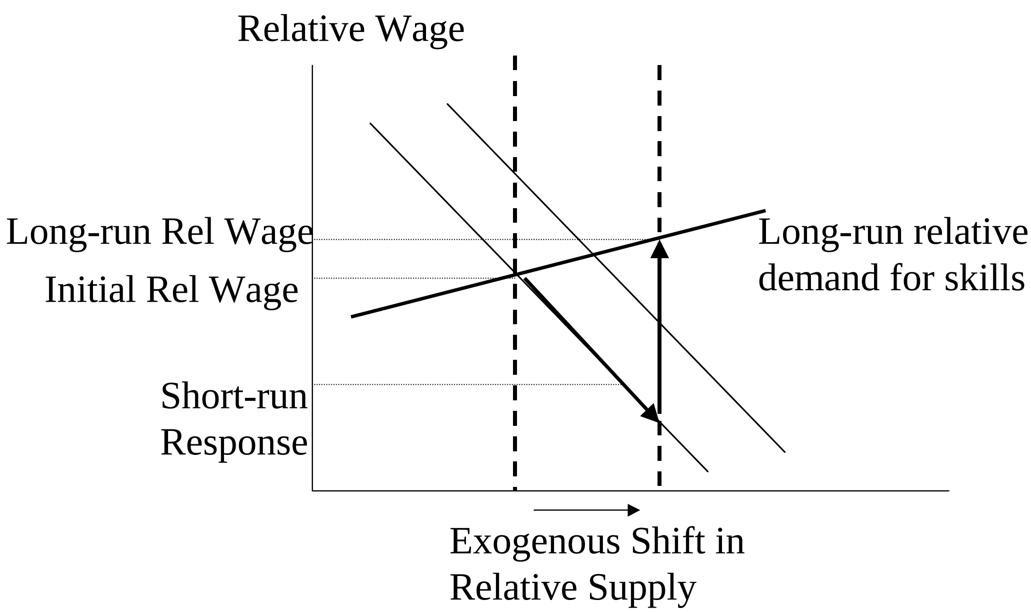
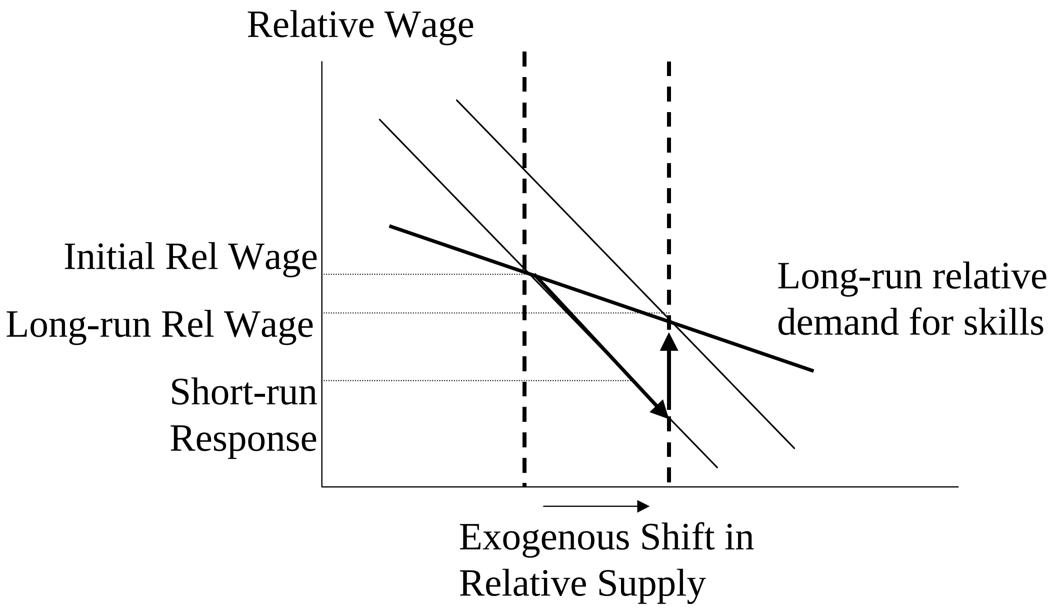

Relative Wage
Long-run Rel Wage
Initial Rel Wage
Short-run
Response
Long-run relative
demand for skills
Exogenous Shift in
Relative Supply

Figure 1:

figure draws this case.



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.

12 Equilibrium Technology Bias: More General Models

The above analysis showed how directed technical change can help us understand changes in the skill bias of technology and thus the changes in the wage structure taking place in the United States and other OECD economies.

However, the model was special as it borrowed the standard endogenous growth structure. It also raises a number of theoretical questions regarding whether results about the direction of technical change derived in this specific model are general or not. I now discuss these issues. For this purpose, I will first introduced three different environments in which technology choice can be analyzed.

12.1 The Basic Environments

Consider an economy consisting of a unique final good and two sets of factors of production, a total of $M+N$, $\mathbf{L}=(L_1,...,L_M)$ and $\mathbf{Z}=(Z_1,...,Z_N)$. Throughout, I assume that all factors are supplied inelastically and denote their supplies by $\bar{\mathbf{L}}$ and $\bar{\mathbf{Z}}$. The reason for distinguishing between these two sets of factors is to carry out comparative static exercises varying the supply of factors $\mathbf{Z}$, while holding the supply of other factors, $\mathbf{L}$, constant. The economy consists of a continuum of firms denoted by the set $\mathcal{F}$, each with an identical production function. Without loss of any generality let us normalize the measure of $\mathcal{F}$, $|\mathcal{F}|$, to 1. The price of the final good is always normalized to 1.

12.1.1 Economy D—Decentralized Equilibrium

In the first environment, *Economy D,* all markets are competitive and technology is decided by each firm separately. In this case, each firm $i\in\mathcal{F}$ has access to a production function

$$Y^i = F(\mathbf{L}^i, \mathbf{Z}^i, \theta^i) \tag{48}$$

where $\mathbf{L}^i \in \mathcal{L} \subset \mathbb{R}^M_+, \mathbf{Z}^i \in \mathcal{Z} \subset \mathbb{R}^N_+$ and $\theta^i \in \Theta$ is the measure of technology. F is a real-valued production function, which, for simplicity, I take to be twice continuously differentiable in $(\mathbf{L}^i, \mathbf{Z}^i)$. For now I impose no structure on the set Θ, but for concreteness, one might think of $\Theta \subset \mathbb{R}^K$ for some $K \in \mathbb{N}$. For many instances of technology choice, Θ may consist of distinct elements, so it may not be a convex set. For the global results, we will need that both Θ and $\mathcal{Z}$ are lattices according to some order.

Each firm maximizes profits, i.e., it solves the problem:

$$\max_{\mathbf{L}^i \in \mathcal{L}, \mathbf{Z}^i \in \mathcal{Z}, \theta_i \in \Theta} \pi(\mathbf{L}^i, \mathbf{Z}^i, \theta^i) = F(\mathbf{L}^i, \mathbf{Z}^i, \theta^i) - \sum_{j=1}^{M} w_{Lj} L^i_j - \sum_{j=1}^{N} w_{Zj} Z^i_j, \tag{49}$$

where w_{Lj} is the price of factor L_j for $j = 1, ..., M$ and w_{Zj} is the price of factor Z_j for $j = 1, ..., N$, all taken as given by the firm. Similar to the notation for L and Z, I will use $\mathbf{w}_L$ and $\mathbf{w}_Z$ to denote the vector of factor prices. Since there is a total supply $\bar{L}_j$ of factor L_j and a total supply $\bar{Z}_j$ of factor Z_j, and both factors are supplied inelastically, market clearing requires

$$\int_{i \in \mathcal{F}} L^i_j \leq \bar{L}_j \text{ for } j = 1, ..., M \text{ and } \int_{i \in \mathcal{F}} Z^i_j \leq \bar{Z}_j \text{ for } j = 1, ..., N. \tag{50}$$

Definition 1 A competitive equilibrium in Economy D is a set of decisions $\left\{\mathbf{L}^i, \mathbf{Z}^i, \theta^i\right\}_{i \in \mathcal{F}}$ and factor prices $(\mathbf{w}_L, \mathbf{w}_Z)$ such that $\left\{\mathbf{L}^i, \mathbf{Z}^i, \theta^i\right\}_{i \in \mathcal{F}}$ solve (49) given prices $(\mathbf{w}_L, \mathbf{w}_Z)$ and (50) holds.

I refer to any θ^i that is part of the set of equilibrium allocations, $\left\{\mathbf{L}^i, \mathbf{Z}^i, \theta^i\right\}_{i\in\mathcal{F}}$, as "equilibrium technology".

Assumption 1 $F(\mathbf{L}^i, \mathbf{Z}^i, \theta^i)$ is jointly strictly concave in $(\mathbf{L}^i, \mathbf{Z}^i, \theta^i)$ and increasing in $(\mathbf{L}^i, \mathbf{Z}^i)$ and $\mathcal{L}$, $\mathcal{Z}$ and Θ are convex.

Then by standard arguments we have:

Lemma 1 ***(Symmetry)*** Suppose Assumption 1 holds. Then in any competitive equilibrium, $(\mathbf{L}^i, \mathbf{Z}^i, \theta^i) = (\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta)$ for all $i \in \mathcal{F}$.

Proof. This proposition follows immediately by the strict concavity of $F(\mathbf{L}^i, \mathbf{Z}^i, \theta^i)$, which implies strict concavity of $\pi(\mathbf{L}^i, \mathbf{Z}^i, \theta^i)$. To obtain a contradiction, suppose that two firms, i and i', choose $(\mathbf{L}^i, \mathbf{Z}^i, \theta^i)$ and $(\mathbf{L}^{i'}, \mathbf{Z}^{i'}, \theta^{i'})$, such that $(\mathbf{L}^i, \mathbf{Z}^i, \theta^i) \neq (\mathbf{L}^{i'}, \mathbf{Z}^{i'}, \theta^{i'})$. This is only possible if $\pi(\mathbf{L}^i, \mathbf{Z}^i, \theta^i) = \pi(\mathbf{L}^{i'}, \mathbf{Z}^{i'}, \theta^{i'})$. Now consider the vector $(\mathbf{L}, \mathbf{Z}, \theta) = \lambda(\mathbf{L}^i, \mathbf{Z}^i, \theta^i) + (1-\lambda)(\mathbf{L}^{i'}, \mathbf{Z}^{i'}, \theta^{i'})$ for some $\lambda \in (0,1)$, which is feasible by the convexity of $\mathcal{L}$, $\mathcal{Z}$ and Θ. Strict concavity implies that $\pi(\mathbf{L}, \mathbf{Z}, \theta) > \lambda\pi(\mathbf{L}^i, \mathbf{Z}^i, \theta^i) + (1-\lambda)\pi(\mathbf{L}^{i'}, \mathbf{Z}^{i'}, \theta^{i'})$, hence $\pi(\mathbf{L}, \mathbf{Z}, \theta) > \pi(\mathbf{L}^i, \mathbf{Z}^i, \theta^i) = \pi(\mathbf{L}^{i'}, \mathbf{Z}^{i'}, \theta^{i'})$, delivering a contradiction. Therefore for all $i \in \mathcal{F}$, we have $(\mathbf{L}^i, \mathbf{Z}^i, \theta^i) = (\mathbf{L}, \mathbf{Z}, \theta)$. Since F is increasing in $(\mathbf{L}^i, \mathbf{Z}^i)$, market clearing, (50), and $|\mathcal{F}| = 1$ imply that $(\mathbf{L}, \mathbf{Z}) = (\bar{\mathbf{L}}, \bar{\mathbf{Z}})$, completing the proof. ■

Assumption 1 may be restrictive, however, because it rules out constant returns to scale in $(\mathbf{L}^i, \mathbf{Z}^i, \theta^i)$. Alternatively, we can modify this assumption to allow for constant returns to scale:

Assumption 1' $\Theta \subset \mathbb{R}^K$ for some $K \geq 1$ and $F(\mathbf{L}^i, \mathbf{Z}^i, \theta^i)$ is increasing in $(\mathbf{L}^i, \mathbf{Z}^i)$ and exhibits constant returns to scale in $(\mathbf{L}^i, \mathbf{Z}^i, \theta^i)$ and $(\bar{\mathbf{L}}, \bar{\mathbf{Z}}) \in \mathcal{L} \times \mathcal{Z}$.

Proposition 1 ***(Welfare Theorem D)*** Suppose Assumption 1 or Assumption 1' holds. Then any equilibrium technology θ is a solution to

$$\max_{\theta' \in \Theta} F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta'), \tag{51}$$

and any solution to this problem is an equilibrium technology.

Proof. ($\Longrightarrow$) First suppose Assumption 1 holds. Suppose that $\left\{\mathbf{L}^i, \mathbf{Z}^i, \theta^i\right\}_{i \in \mathcal{F}}$ is a competitive equilibrium. By Lemma 1, $\left\{\mathbf{L}^i, \mathbf{Z}^i, \theta^i\right\}_{i \in \mathcal{F}}$ is such that $\left(\mathbf{L}^i, \mathbf{Z}^i, \theta^i\right) = \left(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta\right)$ for all $i \in \mathcal{F}$. Moreover, by the definition of a competitive equilibrium, there exist $\mathbf{w}_L$ and $\mathbf{w}_H$ such that

$$\left(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta\right) \in \arg \max_{\mathbf{L}^i \in \mathcal{L}, \mathbf{Z}^i \in \mathcal{Z}, \theta_i \in \Theta} F(\mathbf{L}^i, \mathbf{Z}^i, \theta^i) - \sum_{j=1}^{M} w_{Lj} L_j^i - \sum_{j=1}^{N} w_{Zj} Z_j^i. \tag{52}$$

This implies that any equilibrium technology θ satisfies $\theta \in \arg\max_{\theta' \in \Theta} F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta')$. Next, suppose that Assumption 1' holds. In that case, without loss of any generality, we can consider an equilibrium with only one (representative) firm active and employing $(\bar{\mathbf{L}}, \bar{\mathbf{Z}}) \in \mathcal{L} \times \mathcal{Z}$. Consequently, by the definition of a competitive equilibrium (52) holds. Thus the same conclusion follows, completing the proof.

($\Longleftarrow$) First suppose that Assumption 1 holds. Take $\theta \in \arg\max_{\theta' \in \Theta} F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta')$. Consider the factor price vectors $\mathbf{w}_L$ and $\mathbf{w}_H$ such that $w_{Lj} = \partial F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta)/\partial L_j$

and $w_{Zj} = \partial F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta)/\partial Z_j$. The strict concavity of F and (51) imply that at these factor price vectors, $\left(\mathbf{L}^i, \mathbf{Z}^i, \theta^i\right) = \left(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta\right)$ for all $i \in \mathcal{F}$ satisfies the first-order necessary and sufficient conditions for (52), so it is a competitive equilibrium, thus θ is an equilibrium technology. Next, suppose that Assumption 1' holds. Once again, we can consider an equilibrium with only one firm active employing $(\bar{\mathbf{L}}, \bar{\mathbf{Z}}) \in \mathcal{L} \times \mathcal{Z}$, so θ must be such that $\theta \in \arg\max_{\theta' \in \Theta} F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta')$, completing the proof. ■

Proposition 1 is useful since it enables us to focus on a simple maximization problem rather than an equilibrium problem. We next derive a similar maximization problem for Economies C and M, which relaxed the strong (joint) convexity assumptions inherent in Economy D.

12.1.2 Economy C—Centralized Equilibrium

In this economy, there is still a unique final good and each firm has access to the production function

$$Y^i = G(\mathbf{L}^i, \mathbf{Z}^i, \theta^i) \tag{53}$$

which is similar to (48). In particular, we again have $\mathbf{L}^i \in \mathcal{L} \subset \mathbb{R}_+^M, \mathbf{Z}^i \in \mathcal{Z} \subset \mathbb{R}_+^N$ and $\theta^i \in \Theta$ is the measure of technology, and G is again a real-valued production function that is twice continuously differentiable in $(\mathbf{L}^i, \mathbf{Z}^i)$. Moreover, let $\theta_0 \in \Theta$ be such that

$$G(\mathbf{L}^i, \mathbf{Z}^i, \theta_0) \leq G(\mathbf{L}^i, \mathbf{Z}^i, \theta) \text{ for all } \theta \in \Theta \text{ and for all } \mathbf{L}^i \in \mathcal{L} \text{ and } \mathbf{Z}^i \in \mathcal{Z}\ , \tag{54}$$

which implies that θ_0 is the least productive technology available.

Each firm has free access to technology θ_0 and can also be given the rights to use some other technology $\theta \neq \theta_0$ from a centralized (socially-run) research firm. This research firm can create technology θ at cost $C(\theta)$ from the technology menu, with the normalization $C(\theta_0) = 0$. To simplify the analysis, I assume that the research firm can only choose one technology, which might be, for example, because of the necessity of standardization across firms. This assumption rules out a situation in which the research firm creates two different technologies from the menu, say θ_1 and θ_2, and provides one technology to a subset of firms and the other to the rest. This strategy will not be optimal as long as $C(\theta)$ is sufficiently large because it to would lead to duplication of the costs of creating new technologies. Since it is not central to the focus here, rather than deriving the conditions on $C(\theta)$ to rule out this possibility, I simply assume that choosing two separate technologies from the menu is not possible. In line with the emphasis in Arrow (1962) and Romer (1990), once created, this technology can be costlessly made available to any firm.

All factor markets are again competitive. Consequently, given the technology offer of θ of the research firm, the maximization problem of each firm is

$$\max_{\mathbf{L}^i \in \mathcal{L}, \mathbf{Z}^i \in \mathcal{Z}, \theta^i \in \{\theta_0, \theta\}} \pi(\mathbf{L}^i, \mathbf{Z}^i, \theta^i) = G(\mathbf{L}^i, \mathbf{Z}^i, \theta^i) - \sum_{j=1}^{M} w_{Lj} L_j^i - \sum_{j=1}^{N} w_{Zj} Z_j^i. \quad (55)$$

Given (54), each firm will choose $\theta^i = \theta$.

In addition, the objective of the research firm is to maximize total surplus, or total output. Since $\theta^i = \theta$ for all $i \in \mathcal{F}$, this is equivalent to

$$\max_{\theta^i \in \Theta} \Pi(\theta) = \int_0^1 G(\mathbf{L}^i, \mathbf{Z}^i, \theta) di - C(\theta). \tag{56}$$

This immediately leads to a natural definition of equilibrium:

Definition 2 An equilibrium in Economy C is a set of decisions $\{\mathbf{L}^i, \mathbf{Z}^i, \theta^i\}_{i \in \mathcal{F}}$ where $\theta^i \in \{\theta_0, \theta\}$, and factor prices $(\mathbf{w}_L, \mathbf{w}_Z)$ such that $\{\mathbf{L}^i, \mathbf{Z}^i, \theta^i\}_{i \in \mathcal{F}}$ solve (49) given $(\mathbf{w}_L, \mathbf{w}_Z)$, (50) holds, and the technology choice for the research firm θ maximizes (56).

We now impose weaker versions of Assumptions 1 and 1' on G:

Assumption 2 $G(\mathbf{L}^i, \mathbf{Z}^i, \theta^i)$ is jointly strictly concave and increasing in $(\mathbf{L}^i, \mathbf{Z}^i)$ and $\mathcal{L}$ and $\mathcal{Z}$ are convex.

Assumption 2' $G(\mathbf{L}^i, \mathbf{Z}^i, \theta^i)$ is increasing and exhibits constant returns to scale in $(\mathbf{L}^i, \mathbf{Z}^i)$ and $(\bar{\mathbf{L}}, \bar{\mathbf{Z}}) \in \mathcal{L} \times \mathcal{Z}$.

The important difference between Assumptions 1 and 1' versus Assumptions 2 and 2' is that with the latter, $G(\mathbf{L}^i, \mathbf{Z}^i, \theta^i)$ does not need to be jointly concave in $(\mathbf{Z}^i, \theta)$, which will play an important role in the analysis below.

Proposition 2 ***(Welfare Theorem C)*** Suppose Assumption 2 or Assumption 2' holds. Then any equilibrium technology is a solution to

$$\max_{\theta \in \Theta} G(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta) - C(\theta) \tag{57}$$

and any solution to this problem is an equilibrium technology.

Proof. The proof is similar to that of Proposition 1, and follows again by noting that under Assumption 2 the equilibrium will be symmetric, so $(\mathbf{L}^i, \mathbf{Z}^i, \theta) = (\mathbf{L}, \mathbf{Z}, \theta)$, and because G is increasing in $(\mathbf{L}^i, \mathbf{Z}^i)$, market clearing, (50), yields that $(\mathbf{L}, \mathbf{Z}) = (\bar{\mathbf{L}}, \bar{\mathbf{Z}})$, which implies that (56) is identical to (57). When Assumption 2' holds, there are constant returns to scale in $(\mathbf{L}, \mathbf{Z})$, and $(\bar{\mathbf{L}}, \bar{\mathbf{Z}}) \in \mathcal{L} \times \mathcal{Z}$, so we can once again work with a single firm employing $(\bar{\mathbf{L}}, \bar{\mathbf{Z}})$, and the conclusion follows. ■

Defining $F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta) = G(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta) - C(\theta)$, we obtain that technology choice in Economy C can be characterized as maximizing some function $F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta)$ with respect to $\theta \in \Theta$ as in Economy D. The important difference is that while in Economy D $F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta)$ is by assumption jointly concave in $(\bar{\mathbf{Z}}, \theta)$, the same is not true in Economy C. In particular, in this latter economy, $F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta)$ does not need to be globally concave in θ, and even at the equilibrium it may be non-concave in $(\bar{\mathbf{Z}}, \theta)$.

12.1.3 Economy M—Monopoly Equilibrium

Now I briefly discuss an economy that is similar to Economy C, but features a monopolist supplying technologies to the firms. I take the simplest structure to deliver results similar to Propositions 1 and 2.

In the environment here, there is still a unique final good and each firm has access to the production function

$$Y^i = \alpha^{-\alpha} (1-\alpha)^{-1} \left[G(\mathbf{L}^i, \mathbf{Z}^i, \theta^i)\right]^{\alpha} q\left(\theta^i\right)^{1-\alpha} \tag{58}$$

which is similar to (53), except that $G(\mathbf{L}^i, \mathbf{Z}^i, \theta^i)$ is now a subcomponent of the production function, which depends on θ^i, the technology being used by the firm. This subcomponent needs to be combined with an intermediate good embodying technology θ^i, denoted by $q\left(\theta^i\right)$—conditioned on θ^i to emphasize it embodies technology θ^i. This intermediate good is supplied by the monopolist. The term $\alpha^{-\alpha}(1-\alpha)^{-1}$ in the front is a convenient normalization. This structure is a slight generalization of the endogenous technology models a la Romer, Grossman and Helpman, Aghion and Howitt discussed previously. As before, I assume that $\mathbf{L}^i \in \mathcal{L} \subset \mathbb{R}_+^M, \mathbf{Z}^i \in \mathcal{Z} \subset \mathbb{R}_+^N$ and G is a real-valued production function that is twice continuously differentiable in $(\mathbf{L}^i, \mathbf{Z}^i)$. There is no longer any need for the freely available technology $\theta_0 \in \Theta$.

The technology monopolist can create technology θ at cost $C(\theta)$ from the technology menu, and again I assume that it can only choose one technology. Once created, the technology monopolist can produce as many units of the intermediate good of type θ (i.e., embodying technology θ) at cost normalized to $1-\alpha$ unit of the final good (this is also a convenient normalization, without any substantive implications). It can then set a (linear) price per unit of this intermediate good of type θ, denoted by χ.

All factor markets are again competitive. Consequently, each firm takes the type of available technology, θ, and the price of the intermediate good embodying this technology, χ, as given and maximizes

$$\max_{\substack{\mathbf{L}^i\in\mathcal{L},\mathbf{Z}^i\in\mathcal{Z},\\ q(\theta)\geq 0}} \pi(\mathbf{L}^i,\mathbf{Z}^i,q(\theta)\mid\theta,\chi)=\alpha^{-\alpha}(1-\alpha)^{-1}\left[G(\mathbf{L}^i,\mathbf{Z}^i,\theta)\right]^{\alpha}q(\theta)^{1-\alpha}-\sum_{j=1}^{M}w_{Lj}L_j^i-\sum_{j=1}^{N}w_{Zj}Z_j^i-\chi q(\tag{59}$$

which gives the following simple inverse demand function for intermediates of type θ as a function of its price, χ, and the factor employment levels of the firm as

$$q^i\left(\theta,\chi,\mathbf{L}^i,\mathbf{Z}^i\right)=\alpha^{-1}G(\mathbf{L}^i,\mathbf{Z}^i,\theta)\chi^{-1/\alpha}. \tag{60}$$

The problem of the monopolist is to maximize its profits, which are naturally given by price minus marginal cost of production times total sales of the intermediates, minus the cost of creating the technology. Thus the problem of the monopolist is:

$$\max_{\theta,\chi,[q^i(\theta,\chi,\mathbf{L}^i,\mathbf{Z}^i)]_{i\in\mathcal{F}}} \Pi=(\chi-(1-\alpha))\int_{i\in\mathcal{F}}q^i\left(\theta,\chi,\mathbf{L}^i,\mathbf{Z}^i\right)di-C(\theta) \tag{61}$$

subject to (60). Therefore, an equilibrium in this economy can be defined as:

Definition 3 An equilibrium in Economy M is a set of decisions $\{\mathbf{L}^i,\mathbf{Z}^i,q^i(\theta,\chi,\mathbf{L}^i,\mathbf{Z}^i)\}_{i\in\mathcal{F}}$ and factor prices $(\mathbf{w}_L,\mathbf{w}_Z,\chi)$ such that $\{\mathbf{L}^i,\mathbf{Z}^i,q^i(\theta,\chi,\mathbf{L}^i,\mathbf{Z}^i)\}_{i\in\mathcal{F}}$ solve (59) given $(\mathbf{w}_L,\mathbf{w}_Z,\chi)$, (50) holds, and the technology choice and pricing decision (θ,χ) for the monopolist maximize (61) subject to (60).

The equilibrium in this economy is straightforward to characterize because (60) defines a constant elasticity demand curve, so the optimal price of the monopolist that maximizes (61) is simply the standard monopoly markup, which means $1/(1-\alpha)$ times the marginal cost of production of the intermediate, $1-\alpha$. This leads to an equilibrium monopoly price of $\chi = 1$. Moreover, I continue to impose Assumption 2 or 2', which imply that the equilibrium will be symmetric, so $q^i(\theta, \chi) = \alpha^{-1} G(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta) \chi^{-1/\alpha}$ for all $i \in \mathcal{F}$, and given the monopoly price $\chi = 1$, we have $q^i(\theta) = q^i\left(\theta, \chi = 1, \bar{\mathbf{L}}, \bar{\mathbf{Z}}\right) = G(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta)$ for all $i \in \mathcal{F}$. The profits and the maximization problem of the monopolist can then be expressed as

$$\max_{\theta \in \Theta} \Pi(\theta) = G(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta) - C(\theta). \tag{62}$$

Thus we have established (proof omitted):

Proposition 3 ***(Equilibrium Theorem M)*** Suppose Assumption 2 or Assumption 2' holds. Then an equilibrium in Economy M is a solution to

$$\max_{\theta \in \Theta} G(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta) - C(\theta)$$

and any solution to this problem is an equilibrium.

I refer to this proposition as "equilibrium theorem," since in contrast to Economies D and C, the presence of the monopoly markup implies that the equilibrium is no longer the optimal allocation. More important for our purposes here, however, is that again defining $F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta) = G(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta) - C(\theta)$, equilibrium technology in Economy M is a solution to a problem identical to that in Economy C, and quite similar to the one in Economy D. As in Economy C, $F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta)$ need not be globally concave in θ nor even locally concave in $(\bar{\mathbf{Z}}, \theta)$ in the neighborhood of the equilibrium.

This result therefore shows that for the analysis of equilibrium bias, it is not important whether technology choices are at the firm level or at the centralized level (resulting from some R&D or other research process), and also whether they are made to maximize social surplus or monopoly profits.

12.2 Relative Equilibrium Bias

We have seen that in three different environments, with different market structures and conceptions of technology choice, the characterization of equilibrium technology boils down to an essentially identical maximization problem, i.e., the maximization of some function $F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta)$ where $\bar{\mathbf{L}}$ and $\bar{\mathbf{Z}}$ are the factor supplies in the economy.

Now I make use of this characterization to derive a number of results about equilibrium bias of technology choice. This section analyzes relative equilibrium bias, and for that reason, I focus on a more specialized economy with only two factors, L and Z (i.e., $M = 1$ and $N = 1$). Recall that, in a two-factor economy, *relative equilibrium bias* is defined as the effect of technology on the marginal product (price) of a factor relative to the marginal product (price) of the other factor. More formally, for this section, suppose that F is twice differentiable in $\theta \in \Theta$ and as usual in $Z \in \mathcal{Z}$ and $L \in \mathcal{L}$. Denote the marginal product (or price) of the two factors by

$$w_Z(Z, L, \theta) = \partial F(Z, L, \theta) / \partial Z$$

and

$$w_L(Z, L, \theta) = \partial F(Z, L, \theta) / \partial L$$

when their employment levels are given by (Z, L) and the technology is θ.

Recall that $F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta)$ either corresponds to the production function of the firms (Economy D) or we have $F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta) = G(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta) - C(\theta)$, where $G(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta)$ is the production function of the firms (Economy C) or a subcomponent of the production function (Economy M). In both cases, the derivatives of F with respect to $\mathbf{Z}$ and $\mathbf{L}$ define the marginal products of these factors. With a slight abuse of terminology, I will refer to $F(\bar{\mathbf{L}}, \bar{\mathbf{Z}}, \theta)$ as "the production function". From the twice differentiability of F, these marginal products are also differentiable functions of Z and L.

Finally, let Θ be a convex compact subset of $\mathbb{R}^K$ for some $K \geq 1$ and denote the jth component of $\theta \in \Theta$ by θ_j. Then we have the following definitions:

Definition 4 Let $\theta \in \Theta \subset \mathbb{R}^K$. An increase in technology θ_j for some $j = 1, ..., K$ is *relatively biased* towards factor Z at $\left(\bar{Z}, \bar{L}, \theta\right) \in \mathcal{Z} \times \mathcal{L} \times \Theta$ if

$$\frac{\partial w_Z\left(\bar{Z}, \bar{L}, \theta\right) / w_L\left(\bar{Z}, \bar{L}, \theta\right)}{\partial \theta_j} \geq 0.$$

This definition simply expresses what it means for a technology to be relatively biased towards a factor (similarly a decrease in θ_j is relatively biased towards factor Z, if the derivative in Definition 4 is non-positive). From this definition, it is clear that (weak) relative equilibrium bias should correspond to a change in technology θ in a direction biased towards Z *in response to an increase* in $\bar{Z}$ (or $\bar{Z}/\bar{L}$); this is stated in the second definition.

Definition 5 Let $\theta \in \Theta \subset \mathbb{R}^K$, denote the equilibrium technology at factor supplies $(\bar{Z}, \bar{L}) \in \mathcal{Z} \times \mathcal{L}$ by $\theta(\bar{Z}, \bar{L})$, and assume that $d\theta_j(\bar{Z}, \bar{L})/dZ/L$ exists at $(\bar{Z}, \bar{L})$ for all for all $j = 1, ..., K$. Then there is *relative equilibrium bias* at $(\bar{Z}, \bar{L}, \theta(\bar{Z}, \bar{L}))$ if

$$\sum_{j=1}^{K} \frac{\partial w_Z(\bar{Z}, \bar{L}, \theta(\bar{Z}, \bar{L}))/w_L(\bar{Z}, \bar{L}, \theta(\bar{Z}, \bar{L}))}{\partial \theta_j} \frac{d\theta_j(\bar{Z}, \bar{L})}{dZ/L} \geq 0. \tag{63}$$

Notice that the definition of relative equilibrium bias requires the (overall) change in technology in response to an increase in $\bar{Z}/\bar{L}$ to be biased towards Z at some point $(\bar{Z}, \bar{L}) \in \mathcal{Z} \times \mathcal{L}$ for which $d\theta_j(\bar{Z}, \bar{L})/d(Z/L)$ exists. This requirement entails two restrictions. The first is the usual non-singularity requirement to enable an application of the implicit function theorem, i.e., that the Hessian of F with respect to θ, $\nabla^2_{\theta\theta} F(\bar{Z}, \bar{L}, \theta(\bar{Z}, \bar{L}))$, is non-singular. The second is more subtle; since we have not made global concavity assumptions (except in Economy D), a small change in Z may shift the technology choice from one local optimum to another, thus essentially making $d\theta_j(\bar{Z}, \bar{L})/d(Z/L)$ infinite (or undefined). This possibility is also ruled out by this assumption. This discussion suggests that the assumption that $d\theta_j(\bar{Z}, \bar{L})/d(Z/L)$ exists at $(\bar{Z}, \bar{L})$ can be replaced by the following:

Assumption A1: $\nabla^2_{\theta\theta} F(\bar{Z}, \bar{L}, \theta(\bar{Z}, \bar{L}))$ is non-singular and there exists $\delta > 0$ such that for all $\theta' \in \Theta$ with $\partial F(\bar{Z}, \bar{L}, \theta')/\partial\theta = 0$, we have $F(\bar{Z}, \bar{L}, \theta(\bar{Z}, \bar{L})) - F(\bar{Z}, \bar{L}, \theta') > \delta$.

The second part of the assumption ensures that the peaks of the function $F(\bar{Z}, \bar{L}, \theta)$ in θ are "well separated", in the sense that in response to a small change in factor supplies, there will not be a shift in the global optimum of θ from one local optimum to another. Consequently, Assumption A1 is equivalent to assuming that $d\theta_j\left(\bar{Z},\bar{L}\right)/d\left(Z/L\right)$ or $d\theta_j\left(\bar{Z},\bar{L}\right)/dZ$ exists at $\left(\bar{Z},\bar{L}\right)$ for all j. Nevertheless, it is more intuitive (and straightforward) to directly assume that these derivatives exist rather than imposing Assumption A1, but depending on taste, this assumption can be substituted throughout.

The next definition introduces the more stringent concept of strong relative bias, which requires that in response to an increase in $\bar{Z}$ (or $\bar{Z}/\bar{L}$), technology changes so much that the overall effect (after the induced change in technology) is to increase the relative price of factor Z. For this definition, recall that $\nabla^2_{xy}F$ denotes the Hessian of F with respect to x and y.

Definition 6 Let $\theta \in \Theta \subset \mathbb{R}^K$, denote the equilibrium technology at factor supplies $\left(\bar{Z},\bar{L}\right) \in \mathcal{Z}\times\mathcal{L}$ by $\theta\left(\bar{Z},\bar{L}\right)$, and assume that $d\theta_j\left(\bar{Z},\bar{L}\right)/dZ/L$ exists at $\left(\bar{Z},\bar{L}\right)$ for all $j=1,...,K$. Then there is *strong relative equilibrium bias* at $\left(\bar{Z},\bar{L},\theta\left(\bar{Z},\bar{L}\right)\right)$ if

$$\left.\frac{\partial w_Z\left(\bar{Z},\bar{L},\theta\left(\bar{Z},\bar{L}\right)\right)/w_L\left(\bar{Z},\bar{L},\theta\left(\bar{Z},\bar{L}\right)\right)}{\partial Z/L}\right|_{\theta}+\sum_{j=1}^{K}\frac{\partial w_Z\left(\bar{Z},\bar{L},\theta\left(\bar{Z},\bar{L}\right)\right)/w_L\left(\bar{Z},\bar{L},\theta\left(\bar{Z},\bar{L}\right)\right)}{\partial\theta_j}\frac{d\theta_j}{dZ/L}> \tag{64}$$

By comparing the latter two definitions, it is clear that there will be strong relative equilibrium bias if the sum of the expressions in (63) over $j=1,...,K$ is large enough to dominate the direct (negative) effect of the increase in relative supplies on relative wages (which is the first term in (64)).

The main result in this section is that the conjecture about relative equilibrium bias applies in a world with only factor-augmenting technologies, but not more generally. Before deriving these results, it is useful to return to the basic results derived above for the endogenous growth models, but rederived them now in the context of Economy C or M. In particular, the next example considers an environment equivalent to Economy C or M above, with constant returns to scale in L and Z.

Example 1 ***(Relative Equilibrium Bias)*** Suppose that output (i.e., $G(\mathbf{L}^i, \mathbf{Z}^i, \theta)$) is given by a constant elasticity of substitution aggregate production function

$$Y = \left[\gamma \left(A_L L\right)^{\frac{\sigma-1}{\sigma}} + (1-\gamma)\left(A_Z Z\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}, \tag{65}$$

where A_L and A_Z are two separate factor augmenting technology terms, $\gamma \in (0,1)$ is a distribution parameter which determines how important the two factors are, and $\sigma \in (0,\infty)$ is the elasticity of substitution between the two factors. When $\sigma = \infty$, the two factors are perfect substitutes, and the production function is linear. When $\sigma = 1$, the production function is Cobb-Douglas, and when $\sigma = 0$, there is no substitution between the two factors, and the production function is Leontieff. Since there are two technology terms, I take $\theta = (A_L, A_Z) \in \Theta = \mathbb{R}^2_+$.

Suppose that factor supplies are given by $\left(\bar{Z}, \bar{L}\right)$. Then the relative marginal product of the two factors is:

$$\frac{w_Z}{w_L} = \frac{1-\gamma}{\gamma}\left(\frac{A_Z}{A_L}\right)^{\frac{\sigma-1}{\sigma}}\left(\frac{\bar{Z}}{\bar{L}}\right)^{-\frac{1}{\sigma}}. \tag{66}$$

The relative marginal product of Z is decreasing in the relative abundance of Z, $\bar{Z}/\bar{L}$. This is the usual substitution effect, leading to a downward-sloping relative demand curve. This expression also makes it clear that the measure of relative bias towards Z should be defined as $\bar{\theta} = (A_Z/A_L)^{(\sigma-1)/\sigma}$, Alternatively, we could define $A_Z^{(\sigma-1)/\sigma}$ and $A_L^{-(\sigma-1)/\sigma}$ as two separate technology terms, both relatively biased towards Z, but clearly focusing on $\bar{\theta}$ is more economical.

It is important that the bias towards factor Z is $\bar{\theta} = (A_Z/A_L)^{(\sigma-1)/\sigma}$, not A_Z/A_L, as is sometimes confusingly and incorrectly stated in the applied literature. A_Z/A_L is the ratio of Z-augmenting to L-augmenting technology. When $\sigma > 1$, an increase in A_Z/ A_L increases the relative marginal product of Z, while when $\sigma < 1$, an increase in A_Z/A_L reduces the relative marginal product of Z.

This example can also be used to define "absolute bias," but I leave this to the next section. since higher levels of $\bar{\theta}$ increase the marginal product of Z relative to labor for all values of σ. To derive the results similar to those in Acemoglu (1998, 2002) in the context of Economy C or M, suppose that θ_0 corresponds to $A_Z = 0$ and $A_L = 0$, and assume that the costs of producing new technologies are $\eta_L A_L^{1+\delta}$ and $\eta_Z A_Z^{1+\delta}$, where $\delta > 0$. Despite the fact that $\delta > 0$ introduces diminishing returns in the choice of technology, the production possibilities set of this economy is non-convex, since there is choice both over the factors of production, L and Z, and the technologies, A_L and A_Z (so that the function (65) exhibits increasing returns in L, Z,

A_L and A_Z). From Proposition 2 or 3, the problem of choosing equilibrium technology now boils down to:

$$\max_{A_L, A_Z} \left[\gamma \left(A_L \bar{L}\right)^{\frac{\sigma-1}{\sigma}} + (1-\gamma)\left(A_Z \bar{Z}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} - \eta_L A_L^{1+\delta} - \eta_Z A_Z^{1+\delta}.$$

Taking the ratio of the first-order conditions with respect to A_L and A_Z, the solution to this problem immediately yields

$$\frac{A_Z}{A_L} = \left(\frac{\eta_Z}{\eta_L}\right)^{-\frac{\sigma}{1+\sigma\delta}} \left(\frac{1-\gamma}{\gamma}\right)^{\frac{\sigma}{1+\sigma\delta}} \left(\frac{\bar{Z}}{\bar{L}}\right)^{\frac{\sigma-1}{1+\sigma\delta}}.$$

This equation can also be expressed in an alternative form, particularly useful for Theorem 2 below. It states that

$$\frac{d\ln\left(A_Z/A_L\right)}{d\ln\left(Z/L\right)} = \frac{\sigma-1}{1+\sigma\delta}. \tag{67}$$

The measure of relative bias towards factor Z is then given as

$$\bar{\theta} = \left(\frac{\eta_Z}{\eta_L}\right)^{-\frac{\sigma-1}{1+\sigma\delta}} \left(\frac{1-\gamma}{\gamma}\right)^{\frac{\sigma-1}{1+\sigma\delta}} \left(\frac{\bar{Z}}{\bar{L}}\right)^{\frac{(\sigma-1)^2}{(1+\sigma\delta)\sigma}}.$$

Since $(\sigma-1)^2/(1+\sigma\delta)\sigma > 0$, this establishes that when $\bar{Z}/\bar{L}$ increases, the relative bias towards factor Z increases for any value of $\sigma \neq 1$ (and remains constant when $\sigma = 1$), which is the basis of the conjecture in the Introduction. More explicitly, returning to the discussion above, when $\sigma > 1$, an increase in $\bar{Z}/\bar{L}$ increases A_Z/A_L, which in turn raises w_Z/w_L at given factor proportions. In contrast when $\sigma < 1$, an increase in $\bar{Z}/\bar{L}$ reduces A_Z/A_L, but in this case, A_Z/A_L is biased against factor Z (it is biased towards factor L), so a decrease in A_Z/A_L again raises w_Z/w_L.

Moreover, substituting this into (66), we obtain

$$\frac{w_Z}{w_L} = \left(\frac{\eta_Z}{\eta_L}\right)^{-\frac{\sigma-1}{1+\sigma\delta}} \left(\frac{1-\gamma}{\gamma}\right)^{\frac{\sigma+\sigma\delta}{1+\sigma\delta}} \left(\frac{\bar{Z}}{\bar{L}}\right)^{\frac{\sigma-2-\delta}{1+\sigma\delta}}, \tag{68}$$

or

$$\frac{d\ln\left(w_Z/w_L\right)}{d\ln\left(Z/L\right)} = \frac{\sigma-2-\delta}{1+\sigma\delta}.$$

so that when $\sigma > 2+\delta$, the relative demand curve for factors is *upward-sloping,* corresponding to the "strong relative endogenous bias" result.

This example therefore illustrates both the possibility of weak and strong relative bias results in an economy within non-convex aggregate production possibilities set. In particular, technological change induced in response to an increase in Z is always (weakly) relatively biased towards Z, and moreover, if the condition $\sigma > 2+\delta$ is satisfied, there is strong relative bias. Nevertheless, the structure of the economy is very specialized, in particular, it incorporates a specific aggregate production function and cost functions for undertaking research. But more important is the assumption that all technologies are assumed to be of the factor-augmenting form. I next establish that with a more general setup, but still with two-factors and factor-augmenting technologies, the same results hold.

Theorem 2 ***(Relative Equilibrium Bias with Factor-Augmenting Technologies)*** Consider Economy C or Economy M with two-factors and factor-augmenting technologies, i.e., a continuously differentiable production function $F(A_L L, A_Z L)$, and costs of producing technologies $C_Z(A_Z)$ and $C_L(A_L)$, which are also assumed to be continuously differentiable, increasing and convex. Let σ be the (local) elasticity of substitution between L and Z, defined by $\sigma = -\left.\frac{\partial \ln(Z/L)}{\partial \ln(w_Z/w_L)}\right|_{\frac{A_Z}{A_L}}$, let $\delta = \frac{\partial \ln\left(C_Z'(A_Z)/C_L'(A_L)\right)}{\partial \ln(A_Z/A_L)}$, and suppose that factor supplies are given by $\left(\bar{Z}, \bar{L}\right)$. Then we have that for all $\left(\bar{Z}, \bar{L}\right) \in \mathcal{Z} \times \mathcal{L}$:

$$\frac{d \ln (A_Z/A_L)}{d \ln (Z/L)} = \frac{\sigma - 1}{1 + \sigma\delta} \tag{69}$$

and

$$\frac{d \ln \left(w_Z \left(A_L \bar{L}, A_Z \bar{Z}\right) / w_L \left(A_L \bar{L}, A_Z \bar{Z}\right)\right)}{d \ln (Z/L)} = \frac{\sigma - 2 - \delta}{1 + \sigma\delta}, \tag{70}$$

so that there is always (weak) *relative equilibrium bias* and there is *strong relative equilibrium bias* if $\sigma - 2 - \delta > 0$.

Proof. From Proposition 2 or 3, we need to look at the following maximization problem:

$$\max_{A_L, A_Z} F\left(A_L \bar{L}, A_Z \bar{Z}\right) - C_Z (A_Z) - C_L (A_L) .$$

Taking the ratio of the first-order conditions gives

$$\frac{\bar{Z}}{\bar{L}} \frac{F_Z \left(A_L \bar{L}, A_Z \bar{Z}\right)}{F_L \left(A_L \bar{L}, A_Z \bar{Z}\right)} = \frac{C_Z' (A_Z)}{C_L' (A_L)}$$

where F_L denotes the derivative of F with respect to its first argument and F_Z is the derivative with respect to the second. Recalling the definition of marginal products, this gives

$$\frac{\bar{Z}}{\bar{L}} \frac{w_Z (A_L L, A_Z Z)}{w_L (A_L L, A_Z Z)} = \frac{\bar{Z}}{\bar{L}} \frac{A_Z}{A_L} \frac{F_Z (A_L L, A_Z Z)}{F_L (A_L L, A_Z Z)} = \frac{A_Z}{A_L} \frac{C_Z' (A_Z)}{C_L' (A_L)}. \tag{71}$$

Now taking logs and differentiating totally with respect to $\ln(Z/L)$ gives:

$$\left(1+\frac{\partial \ln\left(C_Z'(A_Z)/C_L'(A_L)\right)}{\partial \ln(A_Z/A_L)}\right)\frac{d\ln(A_Z/A_L)}{d\ln(Z/L)} = \left.\frac{\partial \ln(w_Z/w_L)}{\partial \ln(Z/L)}\right|_{\frac{A_Z}{A_L}}+1$$
$$+\left.\frac{\partial \ln(w_Z/w_L)}{\partial \ln(A_Z/A_L)}\right|_{\frac{\bar{Z}}{\bar{L}}}\frac{d\ln(A_Z/A_L)}{d\ln(Z/L)}.$$

Since from (71), $\left.\frac{\partial \ln(w_Z/w_L)}{\partial \ln(A_Z/A_L)}\right|_{\frac{\bar{Z}}{\bar{L}}}=\frac{\sigma-1}{\sigma}$, rearranging and recalling the definitions of δ and σ, we obtain

$$\frac{d\ln(A_Z/A_L)}{d\ln(Z/L)}=\frac{\sigma-1}{1+\sigma\delta}$$

as in (69). The result in (70) follows immediately by noting that

$$\frac{d\ln(w_Z/w_L)}{d\ln(Z/L)}=-\frac{1}{\sigma}+\frac{\sigma-1}{\sigma}\frac{d\ln(A_Z/A_L)}{d\ln(Z/L)}.$$

■

The major result of this theorem is that the insights from Example 1 generalize in a very natural way as long as the potential menu of technological possibilities only consists of factor-augmenting technologies. The only difference is that instead of the parameter δ and the elasticity of substitution σ being constants, now they are potentially functions of A_L, and A_Z (and $\bar{L}$ and $\bar{Z}$). Most importantly, the change in A_Z/A_L (or in $(A_Z/A_L)^{(\sigma-1)/\sigma}$ as in Example 1) induced by an increase in $\bar{Z}$ is always relatively biased towards Z, and there is strong equilibrium relative bias if $\sigma>2+\delta$. Therefore, this theorem establishes that an environment with a menu of technological possibilities featuring only factor-augmenting technologies is sufficient to obtain both a *general weak relative bias theorem*, and the *possibility of strong rela-*

tive bias (when the elasticity of substitution between factors, σ, is sufficiently high and the parameter δ is relatively low).

However, once we depart from the world with only factor-augmenting technologies, it is possible for the supply of factor Z to increase, and in response, technology to change in a direction relatively biased against this factor (i.e., towards factor L), thus disproving the conjecture in the Introduction. This is stated in the next theorem and proved by providing a counterexample.

Theorem 3 With a general menu of technologies, there is not necessarily relative equilibrium bias. That is, suppose that $d\theta_j\left(\bar{Z},\bar{L}\right)/dZ/L$ exists at $\left(\bar{Z},\bar{L}\right)$ for all $j=1,...,K$, then

$$\sum_{j=1}^{K}\frac{\partial w_Z\left(\bar{Z},\bar{L},\theta\left(\bar{Z},\bar{L}\right)\right)/w_L\left(\bar{Z},\bar{L},\theta\left(\bar{Z},\bar{L}\right)\right)}{\partial\theta_j}\frac{d\theta_j\left(\bar{Z},\bar{L}\right)}{dZ/L}<0$$

is possible.

Example 4 ***(Counterexample)*** Consider an example of Economy C or M with the family of production functions

$$F\left(L,Z,\theta\right)=A\left(\theta\right)+\left[L^{\theta}+Z^{\theta}\right]^{1/\theta} \tag{72}$$

for $\theta\in\Theta=[a,b]$ where $b>a,$ and $A\left(\theta\right)$ concave and twice continuously differentiable over the entire Θ. [This way of writing the function F incorporates the cost of creating the technology, $C\left(\theta\right)$, in $A\left(\theta\right)$, which is a convenient notation I will adopt in other examples as well.] From Proposition 2 or 3, the choice of θ will maximize $F(L, Z, \theta)$. Therefore, at given factor supplies $(\bar{L}, \bar{Z})$, the equilibrium technology choice θ satisfies

$$\begin{aligned}\frac{\partial F\left(\bar{L}, \bar{Z}, \tilde{\theta}\right)}{\partial\theta} &= A'\left(\tilde{\theta}\right) - \frac{1}{\tilde{\theta}^2}\left[\bar{L}^{\tilde{\theta}} + \bar{Z}^{\tilde{\theta}}\right]^{1/\tilde{\theta}} \ln\left[\bar{L}^{\tilde{\theta}} + \bar{Z}^{\tilde{\theta}}\right] \\ + & \quad \frac{1}{\tilde{\theta}}\left(\bar{L}^{\tilde{\theta}} \ln \bar{L} + \bar{Z}^{\tilde{\theta}} \ln \bar{Z}\right)\left[\bar{L}^{\tilde{\theta}} + \bar{Z}^{\tilde{\theta}}\right]^{\left(1-\tilde{\theta}\right)/\tilde{\theta}} = 0,\end{aligned}$$

with $\partial^2 F\left(\bar{L}, \bar{Z}, \tilde{\theta}\right)/\partial\theta^2 < 0$. Now suppose that $\bar{Z}$ (and thus $\bar{Z}/\bar{L}$) increases. By differentiating this expression with respect to $\bar{Z}$, the effect of this change in $\bar{Z}$ on $\tilde{\theta}$ can be obtained from

$$\begin{aligned}\frac{\partial^2 F\left(\bar{L}, \bar{Z}, \tilde{\theta}\right)}{\partial\theta\partial Z} &= -\frac{1}{\tilde{\theta}^2}\bar{Z}^{\tilde{\theta}-1}\left[\bar{L}^{\tilde{\theta}} + \bar{Z}^{\tilde{\theta}}\right]^{\left(1-\tilde{\theta}\right)/\tilde{\theta}} \ln\left[\bar{L}^{\tilde{\theta}} + \bar{Z}^{\tilde{\theta}}\right] \\ &\quad -\frac{1}{\tilde{\theta}}\bar{Z}^{\tilde{\theta}-1}\left[\bar{L}^{\tilde{\theta}} + \bar{Z}^{\tilde{\theta}}\right]^{\left(1-\tilde{\theta}\right)/\tilde{\theta}} \\ &\quad +\frac{1-\tilde{\theta}}{\tilde{\theta}}\bar{Z}^{\tilde{\theta}-1}\left(\bar{L}^{\tilde{\theta}} \ln \bar{L} + \bar{Z}^{\tilde{\theta}} \ln \bar{Z}\right)\left[\bar{L}^{\tilde{\theta}} + \bar{Z}^{\tilde{\theta}}\right]^{\left(1-2\tilde{\theta}\right)/\tilde{\theta}} \\ &\quad +\frac{1}{\tilde{\theta}}\bar{Z}^{\tilde{\theta}-1}\left(\tilde{\theta}\ln Z + 1\right)\left[\bar{L}^{\tilde{\theta}} + \bar{Z}^{\tilde{\theta}}\right]^{\left(1-\tilde{\theta}\right)/\tilde{\theta}} \\ &\propto -\frac{1}{\tilde{\theta}}\ln\left[\bar{L}^{\tilde{\theta}} + \bar{Z}^{\tilde{\theta}}\right] - 1 \\ &\quad +\left(1-\tilde{\theta}\right)\left(\bar{L}^{\tilde{\theta}} \ln \bar{L} + \bar{Z}^{\tilde{\theta}} \ln \bar{Z}\right)\left[\bar{L}^{\tilde{\theta}} + \bar{Z}^{\tilde{\theta}}\right]^{-1} + \left(\tilde{\theta}\ln \bar{Z} + 1\right)\end{aligned}$$

If this expression is negative, then in response to an increase in $\bar{Z}$ (or $\bar{Z}/\bar{L}$), $\tilde{\theta} = \theta(\bar{L}, \bar{Z})$ will decline. Looking at the ratio of the marginal products of Z and L from (72), we have

$$\frac{w_Z\left(L, Z, \tilde{\theta}\right)}{w_L\left(L, Z, \tilde{\theta}\right)} = \left(\frac{Z}{L}\right)^{\tilde{\theta}-1},$$

which is increasing in $\tilde{\theta}$ as long as $Z > L$. So if $\tilde{\theta}$ declines in response to the increase in $\bar{Z}$ and $\bar{Z} > \bar{L}$, we have a counterexample, showing that in response to an increase in $\bar{Z}$ (or $\bar{Z}/\bar{L}$) θ changes in a direction biased against Z. Such a counterexample is easy to construct. For example, suppose that we start with $\bar{L} = 1$ and $\bar{Z} = 2$, and choose the function $A(\theta)$ such that $\tilde{\theta} = 0.1$. Then

$$\begin{aligned}\frac{\partial^2 F\left(\bar{L}=1, \bar{Z}=2, \tilde{\theta}=0.1\right)}{\partial\theta\partial Z} &\propto -\frac{1}{0.1}\ln\left[1+2^{0.1}\right] - 1 + 0.9\left(2^{0.1}\ln 2\right)\left[1+2^{0.1}\right]^{-1} + (0.1\ln 2 + 1) \\ &\propto -7.28 - 1 + 0.32 + 1.07 < 0,\end{aligned}$$

which is clearly negative. Therefore, in this case the increase in $\bar{Z}$ induces a decline in $\tilde{\theta}$, which is a change in technology relatively biased against Z, providing a counterexample to the conjecture.

Theorem 2 explains the reason for the negative result in Theorem 3. The conjecture about relative bias does not apply in this example because technologies do not take the factor-augmenting form. Although factor augmenting technology may be an interesting and empirically important special case, one may be interested in a more general theorem that applies without imposing a specific structure on the interaction between technologies and factors in the production function. Example 2 shows that such a theorem is not possible for relative bias. In the next section, we will see that such a theorem can be derived for absolute bias. In fact, Example 2 already hints at this possibility. The reason why induced technology (in response to an increase in $\bar{Z}$) is not relatively biased towards Z is that the induced change in technology affects the elasticity of substitution between the two factors, and consequently, even though it increases w_Z (at given factor proportions), it has an even larger (positive) effect on the marginal product of the other factor, w_L.

12.3 Equilibrium Absolute Bias

12.3.1 Local Results

Example 2 shows that there is no general theorem about relative endogenous bias unless we restrict ourselves to factor-augmenting technologies. The obvious question is whether there is a general result for absolute bias. The answer is yes and is the focus of this section. Recall that *absolute bias* refers to whether new technology increases the marginal product of a factor. The main results in this section will therefore show that in response to increases in the supply of a factor (or a set of factors), technology will change endogenously in a direction absolutely biased towards this factor (or this set of factors).

As stated in the Introduction, this section will present both local and global theorems. I begin with the local theorem, which applies to the case with $N = 1$, i.e., to changes in the supply of a single factor, Z.

Given the results in Section 12.1, the problem of equilibrium technology choice is again equivalent to

$$\max_{\theta \in \Theta} F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\right)$$

where $\mathbf{L}$ is a vector of other inputs, $\bar{\mathbf{L}}$ denotes the supply of these other inputs and $\bar{Z}$ denotes the supply of Z. Let us denote the marginal product (or price) of this factor by $w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\right) = \partial F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\right) / \partial Z$ when the employment levels of factors are given by $\left(\bar{Z}, \bar{\mathbf{L}}\right)$ and the technology is θ. For the local result I will also take Θ to be a convex compact subset of $\mathbb{R}^K$ for some $K \geq 1$ and assume that F is also twice differentiable in (Z, θ), which implies that $w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\right)$ is differentiable in θ.

Definition 7 Let $\theta \in \Theta \subset \mathbb{R}^K$. An increase in technology θ_j for some $j = 1, ..., K$ is *locally absolutely biased* towards factor Z at $\left(\bar{Z}, \bar{L}\right) \in \mathcal{Z} \times \mathcal{L}$ if

$$\frac{\partial w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\right)}{\partial \theta_j} \geq 0.$$

Conversely we could define a decrease in technology θ as locally absolutely biased towards factor Z if the same derivative is non-positive. Notice also that the local bias definition requires the bias for only small changes in technology. The global definition below will require a similar directional change but for all magnitudes of changes in supplies. Also this definition can be strengthened to require some strict inequalities. Whether weak or strict inequalities are used is not essential for the essence of the theorem as we will see. Next we define (local) equilibrium absolute bias analogously to relative equilibrium bias.

Definition 8 Let $\theta \in \Theta \subset \mathbb{R}^K$ denote the equilibrium technology at factor supplies $\left(\bar{Z}, \bar{\mathbf{L}}\right) \in \mathcal{Z} \times \mathcal{L}$ by $\theta\left(\bar{Z}, \bar{\mathbf{L}}\right)$ and assume that $d\theta_j\left(\bar{Z}, \bar{\mathbf{L}}\right) / dZ$ exists at $\left(\bar{Z}, \bar{\mathbf{L}}\right)$ for all $j = 1, ..., K$. Then there is *local absolute equilibrium bias* at

$\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)$ if

$$\sum_{j=1}^{K} \frac{\partial w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)}{\partial \theta_j} \frac{d\theta_j\left(\bar{Z}, \bar{\mathbf{L}}\right)}{dZ} \geq 0.$$

In words, this definition requires the induced combined change in the components of technology resulting from an increase in $\bar{Z}$ to move towards increasing the marginal product of factor Z. As in Definition 5 for relative equilibrium bias, this definition also requires $d\theta_j\left(\bar{Z}, \bar{\mathbf{L}}\right)/dZ$ to exist. Once again, the assumption that $d\theta_j\left(\bar{Z}, \bar{L}\right)/dZ$ exists entails the two restrictions mentioned above, and can be alternatively replaced by Assumption A1.

Theorem 5 ***(Local Absolute Bias)*** Suppose that $\Theta \subset \mathbb{R}^K$ and $F\left(Z, \mathbf{L}, \theta\right)$ is twice continuously differentiable in (Z, θ) for all $\theta \in \Theta$, $Z \in \mathcal{Z}$ and $\mathbf{L} \in \mathcal{L}$. Let the equilibrium technology at factor supplies $\left(\bar{Z}, \bar{\mathbf{L}}\right)$ be $\theta\left(\bar{Z}, \bar{\mathbf{L}}\right)$ and assume that $d\theta_j\left(\bar{Z}, \bar{\mathbf{L}}\right)/dZ$ exists at $\left(\bar{Z}, \bar{\mathbf{L}}\right)$ for all $j = 1, ..., K$. Then, there is *local absolute equilibrium bias* for all $\left(\bar{Z}, \bar{\mathbf{L}}\right) \in \mathcal{Z} \times \mathcal{L}$, i.e.,

$$\sum_{j=1}^{K} \frac{\partial w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)}{\partial \theta_j} \frac{d\theta_j\left(\bar{Z}, \bar{\mathbf{L}}\right)}{dZ} \geq 0 \text{ for all } \left(\bar{Z}, \bar{\mathbf{L}}\right) \in \mathcal{Z} \times \mathcal{L}.$$

Moreover, if $d\theta_j\left(\bar{Z}, \bar{\mathbf{L}}\right)/dZ \neq 0$ for some j, then

$$\sum_{j=1}^{K} \frac{\partial w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)}{\partial \theta_j} \frac{d\theta_j\left(\bar{Z}, \bar{\mathbf{L}}\right)}{dZ} > 0.$$

Proof. The proof follows from the implicit function theorem. For expositional clarity, I will first present it for a single technology, θ, and then for the case where all technologies change together. Suppose first that $\theta \in \Theta \in \mathbb{R}$ and that the equilibrium choice of θ is in the interior of Θ, so it must satisfy

$$\frac{\partial F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)}{\partial \theta} = 0, \tag{73}$$

and $\partial^2 F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right) / \partial \theta^2 \leq 0$. Since $d\theta\left(\bar{Z}, \bar{\mathbf{L}}\right) / dZ$ exists at $\left(\bar{Z}, \bar{\mathbf{L}}\right)$, from the implicit function theorem it must be equal to

$$\frac{d\theta\left(\bar{Z}, \bar{\mathbf{L}}\right)}{dZ} = -\frac{\partial^2 F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right) / \partial \theta \partial Z}{\partial^2 F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right) / \partial \theta^2} = -\frac{\partial w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right) / \partial \theta}{\partial^2 F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right) / \partial \theta^2}, \tag{74}$$

so we must have $\partial^2 F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right) / \partial \theta^2 \neq 0$, i.e., $\partial^2 F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right) / \partial \theta^2 < 0$. This in turn implies:

$$\frac{\partial w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)}{\partial \theta} \frac{d\theta\left(\bar{Z}, \bar{\mathbf{L}}\right)}{dZ} = -\frac{\left[\partial w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right) / \partial \theta\right]^2}{\partial^2 F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right) / \partial \theta^2} \geq 0. \tag{75}$$

Moreover, if $d\theta\left(\bar{Z}, \bar{\mathbf{L}}\right) / dZ \neq 0$, then from (74), $\partial w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right) / \partial \theta \neq 0$, so (75) holds with strict inequality. Finally, if θ is not in the interior of Θ, equation (73) may no longer hold, so that $d\theta\left(\bar{Z}, \bar{\mathbf{L}}\right) / dZ = 0$, which then satisfies $\partial w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right) / \partial \theta \times d\theta_j\left(\bar{Z}, \bar{\mathbf{L}}\right) / dZ \geq 0$ with equality.

Next, let us look at all of the changes in technologies together. Let Δw_Z be the change in w_Z

$$\Delta w_Z = \sum_{j=1}^{K} \frac{\partial w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)}{\partial \theta_j} \frac{d\theta_j\left(\bar{Z}, \bar{\mathbf{L}}\right)}{dZ}.$$

Then, we have that

$$\begin{aligned} \Delta w_Z &= \left[\nabla_\theta w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)\right]' \left[\nabla_Z \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right] \\ &= \left[\nabla^2_{\theta Z} F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)\right]' \left[\nabla_Z \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right], \end{aligned} \tag{76}$$

where $\left[\nabla_{\theta} w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)\right]$ is a $K \times 1$ vector of changes in w_Z in response to each component of $\theta \in \Theta \subset \mathbb{R}^K$, and $\left[\nabla_Z \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right]$ is the Jacobian of θ with respect to Z, i.e., a $K \times 1$ vector of changes in each component of θ in response to the change in $\bar{Z}$. The second line uses the fact that w_Z is the derivative of the F function, so $\left[\nabla^2_{\theta Z} F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)\right]$ is also the $K \times 1$ vector of changes in w_Z in response to each component of θ. From the implicit function theorem, we have

$$\nabla_Z \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)' = -\left[\nabla^2_{\theta Z} F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)\right]' \left[\nabla^2_{\theta\theta} F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)\right]^{-1},$$

where $\left[\nabla^2_{\theta\theta} F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)\right]$ is the $K \times K$ Hessian of F with respect to θ, and is symmetric and negative definite by virtue of the fact that $\theta\left(\bar{Z}, \bar{\mathbf{L}}\right)$ is a maximizer according to Propositions 1, 2, or 3, so its inverse, $\left[\nabla^2_{\theta\theta} F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)\right]^{-1}$, is also symmetric and negative definite. Substituting in (76), we obtain

$$\Delta w_Z = -\left[\nabla^2_{\theta Z} F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)\right]' \left[\nabla^2_{\theta\theta} F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)\right]^{-1} \left[\nabla^2_{\theta Z} F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)\right] \geq 0,$$

which establishes the desired result.

It is moreover evident that for the inequality to be strict, it suffices that one element of $\nabla_Z \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)$ be non-zero, i.e., $d\theta_j\left(\bar{Z}, \bar{\mathbf{L}}\right)/dZ \neq 0$ for all j.

The argument for the case where $\theta\left(\bar{Z}, \bar{\mathbf{L}}\right)$ is not in the interior of Θ is identical to above. ■

This theorem therefore shows that once we shift our focus to absolute bias, there is a strong result. Under minimal assumptions, technological change induced by a change in factor supplies will be biased towards the factor that has become more abundant.

There is a clear parallel here with the LeChatelier principle of Samuelson (1947), but also a number of important differences. First, this theorem concerns how marginal products (or prices) change as a result of induced technological changes resulting from changes in factor supplies rather than the elasticity of short-run and long-run demand curves. Second, it applies to the equilibrium of an economy not to the maximization problem of a single firm. Nevertheless, the parallel is also important, since we can think of the change in technology as happening in the "long run", in which case Theorem 5 states that long run changes in marginal products (factor prices) will be less than those in the short run because of induced technological change or technology adoption.

Two shortcomings of Theorem 5 are apparent. First, it applies to changes in the supply of a single factor. Second, it applies only to local (small) changes. Similar to Milgrom and Roberts' (1996) generalization of LeChatelier principle, there is a global version of Theorem 5, and it also uses tools from the theory of monotone comparative statics. I start with changes in a single factor, and then generalize it to multiple factors.

Definition 9 Let $\theta\left(\bar{Z}, \bar{\mathbf{L}}\right)$ be the equilibrium technology choice in an economy with factor supplies $\left(\bar{Z}, \bar{\mathbf{L}}\right)$. We say that there is *global absolute equilibrium* bias if for any $\bar{Z}', \bar{Z} \in \mathcal{Z}$,

$$\bar{Z}' \geq \bar{Z} \Longrightarrow w_Z\left(\tilde{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}', \bar{\mathbf{L}}\right)\right) \geq w_Z\left(\tilde{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right) \text{ for all } \tilde{Z} \in \mathcal{Z} \text{ and } \bar{\mathbf{L}} \in \mathcal{L}.$$

Note that there are two notions of globality in this definition. First, the increase from $\bar{Z}$ to $\bar{Z}'$ is not limited to small changes. Second, the change in technology induced by this increase is required to increase the price of factor Z for all $\tilde{Z} \in \mathcal{Z}$. Once again, this definition can be made stronger by requiring strict inequality.

To state the main results, I need a number of technical definitions.

12.3.2 Some Technical Definitions and Results

In this section, I define some of the terms used in the analysis of global equilibrium bias. The reader is referred to the much more detail discussion in Topkis (1998), and also to Milgrom and Roberts (1990) and Milgrom and Shannon (1994).

Let X be a partially ordered set, with the order denoted by $\geq$ (or $>$). For example, $X = \mathbb{R}^2$ with the order such that $(x_1', x_2') \geq (>) (x_1, x_2)$ only if $x_1' \geq (>) x_1$ and $x_2' \geq (>) x_2$ is a partially ordered set. In contrast, $X = \mathbb{R}$ with the natural order $\geq$ $(>)$ is an ordered set or a chain. Let $x' \vee x$ denote the *join*, or the least upper bound of two elements of a partially ordered set X. For example, when $X = \mathbb{R}^2$, $(x_1', x_2') \vee (x_1, x_2) = (\max\{x_1, x_1'\}, \max\{x_2, x_2'\})$. Similarly, the *meet*, or the greatest lower bound of two elements of a partially ordered set is denoted by $x' \wedge x$, and for the case where $X = \mathbb{R}^2$, $(x_1', x_2') \wedge (x_1, x_2) = (\min\{x_1, x_1'\}, \min\{x_2, x_2'\})$. X or a subset S of X is a *lattice* if it contains the join and the meet of each pair of its elements. A subset X' of X is a *sublattice* of X (i.e., according to the same order over X) if X' contains the joint and the meet of each pair of its own elements.

Let $f : X \to \mathbb{R}$ be a real-valued function and X be a lattice. Then we have a more general definition of supermodularity than the one in the text:

Definition 10 A real-valued function $f(x)$ defined on a (sub)lattice X is supermodular if

$$f(x') + f(x'') \leq f(x' \vee x'') + f(x' \wedge x'') \tag{77}$$

for all x',$x'' \in X$. Moreover, $f(x)$ is strictly supermodular if it satisfies (77) with strict inequality for all unordered x',$x'' \in X$.

When $f(x)$ is twice continuously differentiable over X, this definition is equivalent to the one in the text.

Another useful definition is that of increasing differences, which potentially weakens the supermodularity requirements. Let X and T be partially ordered sets. Then a function $f(x, t)$ defined on a subset S of $X \times T$ has *increasing differences,* if for all $t'' > t$, $f(x, t'') - f(x, t)$ is increasing in x. The notion of increasing differences relaxes some conditions of supermodularity (i.e., it is a weaker concept). Unfortunately, in our application, increasing differences only helps when it is equivalent to supermodularity.

In the text, I also mentioned single crossing property and quasi-supermodularity. These are defined as follows:

Definition 11 A real-valued function $f(x)$ defined on a (sub)lattice X is quasi-supermodular if for all x',$x'' \in X$,

$$\begin{aligned} f(x') &\leq f(x' \vee x'') \Longrightarrow f(x'') \leq f(x' \wedge x''), \text{ and} \\ f(x') &< f(x' \vee x'') \Longrightarrow f(x'') < f(x' \wedge x''). \end{aligned} \tag{78}$$

Definition 12 Let $f(x,t)$ be a real-valued function defined on $X \times T$ where X and T are partially ordered sets. Then $f(x,t)$ satisfies the single crossing property in (x,t) if $x'' > x'$, $t'' > t'$ and $f(x'',t') \geq f(x',t')$ implies that $f(x'',t'') \geq f(x',t'')$ and $f(x'',t') > f(x',t')$ implies that $f(x'',t'') > f(x',t'')$.

Next, it is useful to state part of Lemma 2.6.5 of Topkis (1998), which was invoked in Example 3.

Lemma 2 *1.* If X_1 and X_2 are lattices and X is a sublattice of $X_1 \times X_2$ and $f(x)$ is quasi-supermodular on X, then $f(x)$ has the single crossing property in (x_1,x_2) and (x_2,x_1).

2. If X_1 and X_2 are chains and X is a sublattice of $X_1 \times X_2$ and $f(x)$ has the single crossing property in (x_1,x_2) and (x_2,x_1), then $f(x)$ is quasi-supermodular on X.

Proof. See Topkis (1998). ■

Finally, the key theorem for the analysis is a version of the monotonicity theorem. I state a simplified version of Theorem 2.8.2 of Topkis (1992):

Theorem 6 ***(Monotonicity Theorem)*** Suppose that X and T are lattices and $f(x,t)$ is supermodular in (x,t) on a sublattice S of $X \times T$, then $\arg\max_{x \in S} f(x,t)$ is increasing in t on $\{t : t \in T,\ \arg\max_{x \in S} f(x,t) \text{ is nonempty}\}$.

Proof. See Topkis (1998). ■

Finally, Theorem 2.8.1 of Topkis (1998) is a weaker version of the theorem presented here, since it requires T to be only a partially ordered set and $f(x,t)$ to be supermodular in x and have increasing differences in (x,t). However, in our context, whenever the relevant functions satisfy the weaker conditions of Theorem 2.8.1, they also satisfy the conditions of the monotonicity theorem as stated here.

Definition 13 Let $x = (x_1, ..., x_n)$ be a vector in $X \subset \mathbb{R}^n$, and suppose that the real-valued function $f(x)$ is twice continuously differentiable in x. Then $f(x)$ is supermodular on X if and only if $\partial^2 f(x) / \partial x_i \partial x_{i'} \geq 0$ for all $x \in X$ and for all $i \neq i'$, and strictly supermodular if the inequality is strict everywhere.

12.3.3 Global Equilibrium Bias Results

Theorem 7 ***(Global Equilibrium Bias)*** Suppose that Θ is a lattice, let $\bar{\mathcal{Z}}$ be the convex hull of $\mathcal{Z}$, and suppose $F(Z, \mathbf{L}, \theta)$ is supermodular in (Z, θ) on $\bar{\mathcal{Z}} \times \Theta$ for all $\mathbf{L} \in \mathcal{L}$, then there is *global absolute equilibrium bias*, i.e., for any $\bar{Z}', \bar{Z} \in \mathcal{Z}$, $\bar{Z}' \geq \bar{Z}$ implies

$$\theta\left(\bar{Z}', \bar{\mathbf{L}}\right) \geq \theta\left(\bar{Z}, \bar{\mathbf{L}}\right) \text{ for all } \bar{\mathbf{L}} \in \mathcal{L}$$

and

$$w_Z\left(\tilde{Z},\bar{\mathbf{L}},\theta\left(\bar{Z}',\bar{\mathbf{L}}\right)\right)\geq w_Z\left(\tilde{Z},\bar{\mathbf{L}},\theta\left(\bar{Z}',\bar{\mathbf{L}}\right)\right)\text{ for all }\tilde{Z}\in\mathcal{Z}\text{ and }\bar{\mathbf{L}}\in\mathcal{L}.$$

Proof. The proof follows from the application of the monotonicity theorem of Topkis (1998)—see Appendix B. Given the supermodularity of $F(Z,\mathbf{L},\theta)$ on $\bar{\mathcal{Z}}\times\Theta$ and the fact that Θ is a lattice and $\mathcal{Z}$ is a subset of $\mathbb{R}$ therefore also a lattice, the monotonicity theorem implies that $\bar{Z}'\geq\bar{Z}\implies\theta\left(\bar{Z}',\bar{\mathbf{L}}\right)\geq\theta\left(\bar{Z},\bar{\mathbf{L}}\right)$ for all $\bar{\mathbf{L}}\in\mathcal{L}$. Next the supermodularity of $F(Z,\mathbf{L},\theta)$ in (Z,θ) on $\bar{\mathcal{Z}}\times\Theta$ also implies that $\partial F\left(\tilde{Z},\bar{\mathbf{L}},\theta\right)/\partial Z$ is increasing in θ for all $\tilde{Z}\in\left[\bar{Z},\bar{Z}'\right]\subset\bar{\mathcal{Z}}$. Since $w_Z\left(\tilde{Z},\bar{\mathbf{L}},\theta\left(\bar{Z}',\bar{\mathbf{L}}\right)\right)=\partial F\left(\tilde{Z},\bar{\mathbf{L}},\theta\left(\bar{Z}',\bar{\mathbf{L}}\right)\right)/\partial Z$, the conclusion follows. ■

An important feature of this theorem, as opposed to the local theorem, Theorem 5, is that consistent with Definition 9, the induced change in technology does not only increase the marginal product of factor Z (which is becoming more abundant) at the current supply, $\bar{Z}$, but at all points in the set $\mathcal{Z}$. In this sense, Theorem 7 is indeed a global theorem, applying both for large magnitudes of changes and applying to all admissible levels of factor supplies.

An immediate corollary of this theorem is also useful to note.

Corollary 1 Suppose that the hypotheses in Theorem 7 hold. If in addition, $F(Z,\mathbf{L},\theta)$ is strictly supermodular in (Z,θ) on $\bar{\mathcal{Z}}\times\Theta$ for all $\mathbf{L}\in\mathcal{L}$, then whenever $\theta\left(\bar{Z}',\bar{\mathbf{L}}\right)>\theta\left(\bar{Z},\bar{\mathbf{L}}\right)$, we have $w_Z\left(\tilde{Z},\bar{\mathbf{L}},\theta\left(\bar{Z}',\bar{\mathbf{L}}\right)\right)>w_Z\left(\tilde{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)$ for all $\tilde{Z}\in\mathcal{Z}$ and $\bar{\mathbf{L}}\in\mathcal{L}$.

In this theorem, the fact that $\theta\left(\bar{Z}', \bar{\mathbf{L}}\right) \geq \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)$ (say rather than $\theta\left(\bar{Z}', \bar{\mathbf{L}}\right) \leq \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)$) is not particularly important, since the order over the set Θ is not specified. It could be that as $\bar{Z}$ increases some measure of technology t declines. But in this case, this measure would correspond to a type of technology biased *against* factor Z. If so, we can simply change the order over this parameter, e.g., considered changes in $\tilde{t} = -t$ rather than t. In fact, by this method, functions that have everywhere negative cross-partials can often be transformed into supermodular functions.

Another important point is that we now have a global version of Theorem 5, but at the expense of introducing more structure. In particular, in addition to the relatively weak assumptions of requiring $\mathcal{Z}$ and Θ to be lattices, we need F to be supermodular in (Z, θ).

The important point to note is that there are also limits to how much Theorems 5 and 7 can be generalized. *First,* Theorem 5 does not apply for large changes in Z. *Second,* the supermodularity requirement in Theorem 7 cannot be dispensed with, nor could it be replaced by the weaker conditions of single-crossing or quasi-supermodularity of Milgrom and Shannon (1994). *Third*, the assumption that F should be supermodular on the convex hull of $\mathcal{Z}$ cannot be dispensed with either. The following example illustrates all these features by constructing a simple economy which does not satisfy a global version of (absolute) bias because the production function is not supermodular $\mathcal{Z}$ (though it satisfies single crossing and quasi-supermodularity and satisfies supermodularity on a non-convex set $\mathcal{Z}$).

Example 8 ***(No Global Bias without Supermodularity)*** Suppose that $F(Z, \mathbf{L}, \theta) = Z + (Z - Z^2/8)\theta + A(\theta) + B(\mathbf{L})$ and $Z \in \mathcal{Z} = [0, 6]$ and $\Theta = [0, 2]$ so that F is everywhere increasing in Z. Suppose also that $A(\theta)$ is a strictly concave and continuously differentiable real-valued function with the following boundary conditions to ensure interior solutions to the choice of technology: $A'(0) > 0$ and $A'(2) = -\infty$ and $B(\mathbf{L})$ is an increasing function. Clearly, $F(Z, \mathbf{L}, \theta)$ is not supermodular in (Z, θ), since the cross-partial between Z and θ changes sign depending on whether Z is greater than or less than 4.

Consider $\bar{Z} = 1$ and $\bar{Z}' = 5$ as two potential supply levels of factor Z and some $\bar{\mathbf{L}} \in \mathcal{L}$. It can be verified easily that $F\left(1, \bar{\mathbf{L}}, \theta\right) = 1 + 7\theta/8 + A(\theta) + B\left(\bar{\mathbf{L}}\right)$, so that $\theta(1)$ satisfies $A'(\theta(1)) = -7/8$, whereas $F\left(5, \bar{\mathbf{L}}, \theta\right) = 5 + 15\theta/8 + A(\theta) + B\left(\bar{\mathbf{L}}\right)$ so that $\theta(5)$ satisfies $A'(\theta(5)) = -15/8$. The strict concavity of $A(\theta)$ implies that $\theta(5) > \theta(1)$. Moreover, $w_Z(Z, \theta) = 1 + (1 - Z/4)\theta$, so $w_Z(5, \theta(5)) = 1 - \theta(5)/4 < w_Z(5, \theta(1)) = 1 - \theta(1)/4$, contrary to the claim in Theorem 7.

This example can also be used to illustrate that supermodularity cannot be replaced by the weaker single-crossing property, since $F(Z, \mathbf{L}, \theta)$ may satisfy single crossing both in (Z, θ) and (θ, Z). To illustrate this, let us take $\Theta = \{\theta(1), \theta(5)\}$ and suppose that $\theta(1) = 0$ and $\theta(5) = 1$. Let us continue to take $\mathcal{Z} = [0, 6]$. First to check single crossing in (Z, θ), note that since $\theta(1) = 0$, $F\left(\bar{Z}', \bar{\mathbf{L}}, \theta(1)\right) > F\left(\bar{Z}, \bar{\mathbf{L}}, \theta(1)\right)$ whenever $\bar{Z}' > \bar{Z}$. Therefore, we only have to check that $F\left(\bar{Z}', \bar{\mathbf{L}}, \theta(5)\right) > F\left(\bar{Z}, \bar{\mathbf{L}}, \theta(5)\right)$ whenever $\bar{Z}' >$

$\bar{Z}$. This immediately follows from the fact that $\theta(5) = 1$, so that for all $\bar{Z}', \bar{Z} \in \mathcal{Z} = [0,6]$ and $\bar{Z}' > \bar{Z}$, $\bar{Z}' + \left(\bar{Z}' - \left(\bar{Z}'\right)^2/8\right) > \bar{Z} + \left(\bar{Z} - \bar{Z}^2/8\right)$. To establish single crossing in (θ, z), let us take $\theta(1) = 0$ and $\theta(5) = 1$ and also suppose that $A(0) > A(1)$. In that case, single crossing in (θ, Z) requires that whenever $\bar{Z}', \bar{Z} \in \mathcal{Z} = [0,6]$ and $\bar{Z}' > \bar{Z}$, and

$$\bar{Z} + \left(\bar{Z} - \bar{Z}^2/8\right) + A(1) + B\left(\bar{\mathbf{L}}\right) > \bar{Z} + A(0) + B\left(\bar{\mathbf{L}}\right)$$

it must also be the case that

$$\bar{Z}' + \left(\bar{Z}' - \left(\bar{Z}'\right)^2/8\right) + A(1) + B\left(\bar{\mathbf{L}}\right) > \bar{Z}' + A(0) + B\left(\bar{\mathbf{L}}\right).$$

The first inequality implies $\left(\bar{Z} - \bar{Z}^2/8\right) > A(0) - A(1) > 0$, and the second one requires $\left(\bar{Z}' - \left(\bar{Z}'\right)^2/8\right) > 0$, which is always satisfied given the previous inequality and $\bar{Z}' > \bar{Z}$ with $\bar{Z}', \bar{Z} \in \mathcal{Z} = [0,6]$. Since by Lemma 2.6.5 of Topkis (1998), when $\mathcal{Z}$ and Θ are chains, single crossing in (Z, θ) and (θ, Z) implies quasi-supermodularity, this also implies that supermodularity cannot be replaced with quasi-supermodularity.

Finally, this example also shows that the assumption that the function needs to be supermodular on the convex hull of $\mathcal{Z}$ cannot be dispensed with. In particular, if we take $\mathcal{Z} = \{1,5\}$ and $\Theta = \{\theta(1), \theta(5)\}$, it can be verified that the function F here satisfies supermodularity on $\mathcal{Z} \times \Theta$. However, it fails to satisfy supermodularity over its convex hull, $\bar{\mathcal{Z}} = [1,5]$.

To see why it is necessary for F to be supermodular over the convex hull of Z, note that the supermodularity of F implies that for $Z'' > Z'$ and

$\theta'' > \theta'$, we have

$$F(Z'', \theta'', \mathbf{L}) + F(Z', \theta', \mathbf{L}) \geq F(Z'', \theta', \mathbf{L}) + F(Z', \theta'', \mathbf{L}).$$

Now, assuming differentiability and applying the fundamental theorem of calculus twice and using the definition of w_Z, we have

$$\int_{Z'}^{Z''} \int_{\theta'}^{\theta''} \frac{\partial w_Z(Z, \theta, \mathbf{L})}{\partial \theta} d\theta dZ \geq 0.$$

However, this does not guarantee that

$$\int_{\theta'}^{\theta''} \frac{\partial w_Z(Z, \theta, \mathbf{L})}{\partial \theta} d\theta \geq 0$$

for all $Z \in [Z', Z'']$ unless F is supermodular over the convex hull of $\{Z', Z''\}$.

There is a natural generalization of Theorem 7 in which the supplies of a set of factors change simultaneously. This is presented in the next theorem. Let the production function be $F(\mathbf{Z}, \mathbf{L}, \theta)$, where $\mathbf{Z} = (Z_{1,...,} Z_N)$. Define the marginal products in the usual way as

$$w_{Zj} = \frac{\partial F(\mathbf{Z}, \mathbf{L}, \theta)}{\partial Z_j} \text{ for } j = 1, ..., N.$$

The notion of equilibrium bias generalizes naturally.

Definition 14 Let $\bar{\mathbf{Z}} \in \mathcal{Z} \subset \mathbb{R}_+^N$, $\bar{\mathbf{L}} \in \mathcal{L}$ and $\theta(\bar{\mathbf{Z}}, \bar{\mathbf{L}})$ be the equilibrium technology choice in an economy with factor supplies $(\bar{\mathbf{Z}}, \bar{\mathbf{L}})$. We say that there is *global absolute equilibrium bias* if for any $\bar{\mathbf{Z}}', \bar{\mathbf{Z}} \in \mathcal{Z}$, $\bar{\mathbf{Z}}' \geq \bar{\mathbf{Z}}$ implies

$$w_{Zj}\left(\tilde{\mathbf{Z}}, \bar{\mathbf{L}}, \theta(\bar{\mathbf{Z}}', \bar{\mathbf{L}})\right) \geq w_{Zj}\left(\tilde{\mathbf{Z}}, \bar{\mathbf{L}}, \theta(\bar{\mathbf{Z}}, \bar{\mathbf{L}})\right) \text{ for all } \left(\tilde{\mathbf{Z}}, \bar{\mathbf{L}}\right) \in \mathcal{Z} \times \mathcal{L} \text{ and for all } j = 1, ..., K.$$

Once again, this definition can be strengthened by introducing strict inequalities.

Theorem 9 ***(Generalized Global Equilibrium Bias)*** Suppose that $\mathcal{Z}$ and Θ are lattices, let $\bar{\mathcal{Z}}$ be the convex hul of $\mathcal{Z}$, and suppose that $F(\mathbf{Z}, \mathbf{L}, \theta)$ is supermodular in $(\mathbf{Z}, \theta)$ on $\bar{\mathcal{Z}} \times \Theta$ for all $\mathbf{L} \in \mathcal{L}$, then there is *global absolute equilibrium bias*, i.e., for any $\bar{\mathbf{Z}}', \bar{\mathbf{Z}} \in \mathcal{Z}$, $\bar{\mathbf{Z}}' \geq \bar{\mathbf{Z}} \in \mathcal{Z}$ implies

$$\theta\left(\bar{\mathbf{Z}}', \bar{\mathbf{L}}\right) \geq \theta\left(\bar{\mathbf{Z}}, \bar{\mathbf{L}}\right) \text{ for all } \bar{\mathbf{L}} \in \mathcal{L}$$

and

$$w_{Zj}\left(\tilde{\mathbf{Z}}, \bar{\mathbf{L}}, \theta\left(\bar{\mathbf{Z}}', \bar{\mathbf{L}}\right)\right) \geq w_{Zj}\left(\tilde{\mathbf{Z}}, \bar{\mathbf{L}}, \theta\left(\bar{\mathbf{Z}}', \bar{\mathbf{L}}\right)\right) \text{ for all } \left(\tilde{\mathbf{Z}}, \bar{\mathbf{L}}\right) \in \mathcal{Z} \times \mathcal{L} \text{ and for all } j = 1, ..., K.$$

Proof. The proof is analogous to that of Theorem 7 follows from the supermodularity of $F(\mathbf{Z}, \mathbf{L}, \theta)$ in $(\mathbf{Z}, \theta)$ on $\bar{\mathcal{Z}} \times \Theta$. ■

It is clear that a corollary to this theorem, similar to Corollary to Theorem 7 can be stated with strict supermodularity. I omit this to avoid repetition.

The results so far concern what was referred to as "weak" bias in the sense that they compare marginal products at a given level of factor supplies. Example 1 illustrated the possibility of strong (relative) bias where technology might be so responsive to factor supply changes that when a factor becomes more abundant, its relative price and marginal product increases rather than decrease. Although somewhat counterintuitive at first, this is also a possibility in the class of models studied here. But we will see that it requires some type of non-convexity either in the technology set or in the production possibilities set by allowing for a structure similar to that of Economy C or Economy M. First I define strong absolute bias, and to simplify the discussion, from now on, I focus on changes in a single factor:

Definition 15 Suppose that $N = 1$, Let $\theta(Z, \mathbf{L})$ be the equilibrium technology choice in an economy with factor proportions $(Z, \mathbf{L})$. We say that there is *strong absolute equilibrium bias* at $\left(\left\{\bar{Z}, \bar{Z}'\right\}, \bar{\mathbf{L}}\right)$ if for some $\bar{\mathbf{L}} \in \mathcal{L}$ and $\bar{Z}, \bar{Z}' \in \mathcal{Z}$ with $\bar{Z} > \bar{Z}'$, we have

$$w_Z\left(\bar{Z}', \bar{\mathbf{L}}, \theta\left(\bar{Z}', \bar{\mathbf{L}}\right)\right) > w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right).$$

Similarly, suppose that $\Theta \subset R^K$, $w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)$ is differentiable in Z and $d\theta_j\left(\bar{Z}, \bar{\mathbf{L}}\right)/dZ$ exists at $\left(\bar{Z}, \bar{\mathbf{L}}\right)$ for all $j = 1, ..., K$. Then we say that there is *strong absolute equilibrium bias* at $\left(\bar{Z}, \bar{\mathbf{L}}\right) \in \mathcal{Z} \times \mathcal{L}$ if

$$\begin{aligned}\frac{dw_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)}{dZ} &= \left.\frac{\partial w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)}{\partial Z}\right|_{\theta\left(\bar{Z}, \bar{\mathbf{L}}\right)} \\ &\quad - \left[\nabla^2_{\theta Z} F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)\right]' \left[\nabla^2_{\theta\theta} F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)\right]^{-1} \left[\nabla^2_{\theta Z} F\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z},\right.\right.\right.\end{aligned}$$

The next theorem shows that there cannot be strong absolute bias in Economy D if Θ is a convex subset of $\mathbb{R}^K$ and if F is supermodular in (Z, θ).

Theorem 10 ***(No Strong Bias in Economy D)*** Suppose that Θ is a convex subset of $\mathbb{R}^K$, let the equilibrium technology at factor supplies $\left(\bar{Z}, \bar{\mathbf{L}}\right)$ be $\theta\left(\bar{Z}, \bar{\mathbf{L}}\right)$, assume that $d\theta\left(\bar{Z}, \bar{\mathbf{L}}\right)/dZ$ exists at $\left(\bar{Z}, \bar{\mathbf{L}}\right)$. Then there cannot be *strong absolute bias* in Economy D.

Proof. Let us start with the local result and the case with $\theta \in \mathbb{R}$. Let factor supplies be $(\bar{Z}, \bar{\mathbf{L}})$. Strong absolute bias corresponds to

$$\frac{dw_Z(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}}))}{dZ} = \left.\frac{\partial w_Z(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}}))}{\partial Z}\right|_{\theta(\bar{Z}, \bar{\mathbf{L}})} + \frac{\partial w_Z(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}}))}{\partial \theta}\frac{d\theta}{dZ} > 0.$$

This is equivalent to

$$\frac{dw_Z(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}}))}{dZ} = \left.\frac{\partial^2 F(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}}))}{\partial Z^2}\right|_{\theta(\bar{Z}, \bar{\mathbf{L}})} + \frac{\partial w_Z(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}}))}{\partial \theta}\frac{d\theta}{dZ} > 0.$$

Recall from the proof of Theorem 5 that

$$\frac{d\theta}{dZ} = -\frac{\partial^2 F(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}})) / \partial\theta\partial Z}{\partial^2 F(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}})) / \partial\theta^2} = -\frac{\partial w_Z(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}})) / \partial\theta}{\partial^2 F(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}})) / \partial\theta^2}.$$

so

$$\frac{dw_Z(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}}))}{dZ} = \left.\frac{\partial^2 F(\bar{Z}, \bar{\mathbf{L}}, \theta)}{\partial Z^2}\right|_{\theta(\bar{Z}, \bar{\mathbf{L}})} - \frac{\left(\partial^2 F(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}})) / \partial\theta\partial Z\right)^2}{\partial^2 F(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}})) / \partial\theta^2} > 0$$

which is impossible from the joint concavity of $F(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}}))$ in (Z, θ) which implies that

$$\left(\partial^2 F(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}})) / \partial\theta^2\right) \times \left(\partial^2 F(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}})) / \partial Z^2\right) \geq \left(\partial^2 F(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}})) / \partial\theta\partial Z\right)^2$$

and $\partial^2 F(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}})) / \partial\theta^2 < 0$ (given the non-singularity assumption $\partial^2 F(\bar{Z}, \bar{\mathbf{L}}, \theta) / \partial\theta_j^2 \neq 0$).

The result immediately generalizes to multiple dimensions of technology, i.e., to the case with $\theta \in \mathbb{R}^K$ for $K > 1$. In this case, we have

$$\begin{aligned}\frac{dw_Z(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}}))}{dZ} &= \left.\frac{\partial^2 F(\bar{Z}, \bar{\mathbf{L}}, \theta)}{\partial Z^2}\right|_{\theta(\bar{Z}, \bar{\mathbf{L}})} \\ &\quad - \left[\nabla^2_{\theta Z} F(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}}))\right]' \left[\nabla^2_{\theta\theta} F(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z}, \bar{\mathbf{L}}))\right]^{-1} \left[\nabla^2_{\theta Z} F(\bar{Z}, \bar{\mathbf{L}}, \theta(\bar{Z},\end{aligned}$$

which is non-positive from the joint concavity of F in (Z, θ).

Finally, to prove the global result, note that from the fundamental theorem of calculus, for any $\bar{Z}' > \bar{Z}$, we have

$$w_Z\left(\bar{Z}', \bar{\mathbf{L}}, \theta\left(\bar{Z}', \mathbf{L}\right)\right) - w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right) = \int_{\bar{Z}}^{Z'} \frac{dw_Z\left(Z, \bar{\mathbf{L}}, \theta\left(Z, \bar{\mathbf{L}}\right)\right)}{dZ} dZ.$$

Supermodularity rules out changes in the sign of these cross-partials, thus $dw_Z\left(Z, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)/dZ \leq 0$ for all $Z \in \left[\bar{Z}, \bar{Z}'\right]$, the integral is non-positive, establishing the global result. ■

The result in this theorem is not surprising. In Economy D, the production possibilities set is convex, so the marginal product of each factor is decreasing in this supply even after technology adjusts. In contrast, once we allow for non-convexities, the results are very different. To illustrate the importance of non-convex production possibilities set, I will look both at the version of Economy D with a non-convex technology set, or at Economies C or M, which allow for natural non-convexities, strong absolute bias is possible. [Recall that in Economy D, when Assumption 1 applies the technology set Θ is also assumed to be convex. This is assumption can be relaxed without affecting the analysis, and in fact, it is not implied by Assumption 1', which only requires $\mathcal{L}$ and $\mathcal{Z}$ to be convex.]

Theorem 11 ***(Strong Absolute Bias)*** *Strong absolute equilibrium bias* is possible either in Economy D with a non-convex technology set, Θ, or in Economies C or M.

Proof. This theorem will be proved by providing two examples where strong absolute bias takes place. First, take Economy D and suppose that $F(Z, \mathbf{L}, \theta) = Z^{1/2}\theta^{1/2} - \theta + B(\mathbf{L})$ and $\Theta = \{1, 4\}$. Imagine an increase in $\bar{Z}$ from 4 to $9 + \varepsilon$ where $\varepsilon > 0$. It is straightforward to check that for any $\bar{\mathbf{L}} \in \mathcal{L}$, $F\left(4, \bar{\mathbf{L}}, 1\right) = 2 - 1 + B\left(\bar{\mathbf{L}}\right) > F\left(4, \bar{\mathbf{L}}, 4\right) = 4 - 4 + B\left(\bar{\mathbf{L}}\right)$, so $\theta(4) = 1$. In contrast, $F\left(9 + \varepsilon, \bar{\mathbf{L}}, 4\right) = (9 + \varepsilon)^{1/2} 2 - 4 + B\left(\bar{\mathbf{L}}\right) > F\left(9 + \varepsilon, \bar{\mathbf{L}}, 1\right) = (9 + \varepsilon)^{1/2} - 1 + B\left(\bar{\mathbf{L}}\right)$, so that $\theta(9 + \varepsilon) = 4$ (in particular, the two sides are equal when $\varepsilon = 0$, and the left-hand side increases faster in ε). Therefore, an increase in $\bar{Z}$ from 4 to $9 + \varepsilon$ will induce a change in technology from $\theta(4) = 1$ to $\theta(9 + \varepsilon) = 4$. The price (marginal product) of factor Z is given by $w_Z\left(\bar{Z}, \bar{\mathbf{L}}, \theta\right) = \left(\theta/\bar{Z}\right)^{1/2}/2$, so the change in this price resulting from the increase in $\bar{Z}$ (after technology adjusts) is $w_Z\left(\bar{Z} = 9 + \varepsilon, \bar{\mathbf{L}}, 4\right) - w_Z\left(\bar{Z} = 4, \bar{\mathbf{L}}, 1\right) = (4/(9 + \varepsilon))^{1/2}/2 - (1/4)^{1/2}/2 \simeq 1/3 - 1/4 > 0$ for ε sufficiently small, establishing the possibility of strong absolute bias in Economy D with a non-convex technology set.

Next, consider Economy C or M, and to illustrate that a non-convex technology set is not necessary in these economies, take $\Theta = [0, 4]$. Suppose $F(Z, \mathbf{L}, \theta) = Z + Z^{1/2}\theta - 3\theta/2 + B(\mathbf{L})$ (which is not jointly concave in Z and θ). Now consider a change from $\bar{Z} = 1$ to $\bar{Z} = 4$. Clearly $\theta(1) = 0$ while $\theta(4) = 4$. Moreover, for any $\bar{\mathbf{L}} \in \mathcal{L}$, $w_Z\left(\bar{Z} = 1, \bar{\mathbf{L}}, \theta(1)\right) = 1 < w_Z\left(\bar{Z} = 4, \bar{\mathbf{L}}, \theta(4)\right) = 2$, establishing the claim. ■

The importance of this theorem is that, contrary to the standard neoclassical theory, where the increase in the supply of a factor always reduces its price (or marginal product), with endogenous technology choice or technological change, the price of a factor which has become more abundant can increase. This result also distinguishes the approach in this paper from the literature on the LeChatelier principle, which looks at the decision problem of a single firm. As is well-known, the firm's demand curve for a factor is always downward sloping in its own price, so the equilibrium structure (in particular, the equilibrium with aggregate non-convexities) is important for the results in this paper, especially for the possibility of strong equilibrium bias.

Finally, as stated in the Introduction and already hinted in the discussion, "greater non-convexity" makes it more likely that the economy will feature strong absolute bias. This is formalized in the next theorem. Recall that in Economy C or M, $F(Z, \mathbf{L}, \theta) = G(Z, \mathbf{L}, \theta) - C(\theta)$, so marginal product of Z is equivalently given by the derivative of function F or G. Recall also that $F(Z, \mathbf{L}, \theta)$ has to be locally concave in θ for θ to be an equilibrium technology. We say that it is *not jointly concave* in (Z, θ), if its Hessian fails to be negative semi-definite.

Theorem 12 ***(Non-Convexity and Strong Bias)*** Consider Economy C or M. Suppose that Θ is a convex subset of $\mathbb{R}$, let $\theta\left(\bar{Z}, \bar{\mathbf{L}}\right)$ be the equilibrium technology at factor supplies $\left(\bar{Z}, \bar{\mathbf{L}}\right)$ and assume that $d\theta\left(\bar{Z}, \bar{\mathbf{L}}\right)/dZ$ exists at $\left(\bar{Z}, \bar{\mathbf{L}}\right)$. Then there is strong absolute bias at $\left(\bar{Z}, \bar{\mathbf{L}}\right)$ *if and only if* $F(Z, \mathbf{L}, \theta)$ is not jointly concave in (Z, θ) at $\left(\bar{Z}, \bar{\mathbf{L}}, \theta\left(\bar{Z}, \bar{\mathbf{L}}\right)\right)$.

Proof. Recall the proof of Theorem 10, where it was established that for the case of $\theta \in \mathbb{R}$,

$$\frac{dw_Z\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)}{dZ} = \left.\frac{\partial^2 F\left(\bar{Z},\bar{\mathbf{L}},\theta\right)}{\partial Z^2}\right|_{\theta\left(\bar{Z},\bar{\mathbf{L}}\right)} - \frac{\left(\partial^2 F\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)/\partial\theta\partial Z\right)^2}{\partial^2 F\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)/\partial\theta^2}.$$

The fact that $F\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)$ is not jointly concave in (Z,θ) implies that $\left(\partial^2 F\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)/\partial\theta^2\right)\times\left(\partial^2 F\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)/\partial Z^2\right) < \left(\partial^2 F\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)/\partial\theta\partial Z\right)^2$. Since at the optimal technology choice, $\partial^2 F\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)/\partial\theta^2 < 0$, this implies that $dw_Z\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)/dZ > 0$ and there is strong absolute bias at $\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)$ as claimed in the theorem. Conversely, if $F\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)$ is jointly concave in (Z,θ), then $\left(\partial^2 F\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)/\partial\theta^2\right)\times\left(\partial^2 F\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)/\partial Z^2\right) \geq \left(\partial^2 F\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)/\partial\theta\partial Z\right)^2$, which, together with $\partial^2 F\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)/\partial\theta^2 < 0$, implies that $dw_Z\left(\bar{Z},\bar{\mathbf{L}},\theta\left(\bar{Z},\bar{\mathbf{L}}\right)\right)/dZ \leq 0$, completing the proof. ■

This theorem therefore shows that in Economy C or M strong absolute bias will obtain whenever the function $F(Z,\mathbf{L},\theta)$ fails to be jointly concave in (Z,θ), and will not apply whenever it is concave. [Recall that in an equilibrium of Economy C or M, we need $F(Z,\mathbf{L},\theta)$ to be (locally) concave in $(Z,\mathbf{L})$ (from Assumption 2 or 2') and in θ (from optimality), but there is no requirement that it should be jointly concave in (Z,θ).]

This last theorem therefore highlights the importance of non-convexities in generating strong equilibrium bias of technology.

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

A common shortcoming of all the pure technology approaches discussed in the previous chapter is that they do not naturally predict stagnant average wages and/or falling wages for unskilled workers. In the basic framework above, 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.

13.1 Basic Issues

Models of faster technological progress would naturally predict that unskilled workers should benefit from this faster progress. The endogenous technology approach discussed above, 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 (8) from above, in the simplified version of their model discussed above, the wages of unskilled workers is $w_L = \phi_l\,(g)\,a\,[1+\phi_h^\rho (H/L)^\rho]^{(1-\rho)/\rho}$, 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.

13.2 Allocation of Capital

As an alternative, 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. Skilled workers work with the production function

$$Y_h = A_h^\alpha K_h^{1-\alpha} H^\alpha \tag{79}$$

while unskilled workers work with the production function

$$Y_l = A_l^\alpha K_l^{1-\alpha} L^\alpha, \tag{80}$$

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 (80) 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. 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.

Overall, a potential problem for all of these models based on technical change is to account for the decline in the wages of low-skill workers. The effect of technical change on the organization of the labor market may provide an explanation for this decline.

(Chapter head:)Organizational And Institutional Change

13.3 Organizational Change

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.

For example, team production and other high-performance production methods are now widespread in the U.S. economy (e.g., Ichinowski, Prennushi, and Shaw, 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.

Models of organizational change are interesting in part because they often predict a decline in the wages of low-skill workers as a result of organizational change. Moreover, such models can explain both the changes in the organizational work that we observe, and also make some progress towards opening the black-box of "skill-biased technical change).

13.3.1 A simple model

I will first 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

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

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

$$w_H = B_p A_h^\rho h_O^{\rho-1} \left[A_l^\rho L^\rho + A_h^\rho h_O^\rho \right]^{(1-\rho)/\rho} = B_s A_h^\beta, \tag{81}$$

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

$$w_L = B_p A_l^{\rho} L^{\rho-1} \left[A_l^{\rho} L^{\rho} + A_h^{\rho} h_O^{\rho}\right]^{(1-\rho)/\rho} < w_H \tag{82}$$

Now consider an increase in A_h. Differentiating (81) yields $\partial(A_h h_O)/\partial A_h < 0$, which, from (82), 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 discussed above, 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. This is reminiscent to a technological change that makes the new-organization production function more profitable.

13.3.2 Changes in the organization of production with frictions

Now consider a somewhat more structured model which also tries to get to the same issues. The basic idea is that when either the productivity gap between skilled and unskilled workers is limited or when the number of skilled workers in the labor force is small, it will be profitable for firms to create jobs that to employ both skilled and unskilled workers. But when the productivity gap is large or that are a sufficient number of skilled workers, it may become profitable for (some) firms to target skilled workers, designing the jobs specifically for these workers. Then these firms will wait for the skilled workers, and will try to screen the more skill once among the applicants. In the meantime, there will be lower-quality (low capital) jobs specifically targeted at the unskilled.

Suppose that there are two types of workers. The unskilled have human capital (productivity) 1, while the skilled have human capital $\eta > 1$. Denote the fraction of skilled workers in the labor force by ϕ.

Firms choose the capital stock k before they meet a worker, and matching is assumed to be random, in the sense that each firm, irrespective of its physical capital, has exactly the same probability of meeting different types of workers. Once the firm and the worker match, separating is costly, so there is a quasi-rent to be divided between the pair. Here, the economy is assumed to last for one period, so if the firm and worker do not agree they lose all of the output (see Acemoglu, 1999, for the model where the economy is infinite-horizon and agents who do not agree with their partners can resample). Therefore, bargaining will result in workers receiving a certain fraction of output, which I denote by β.

The production function of a pair of worker and firm is

$$y = k^{1-\alpha} h^{\alpha},$$

where k is the physical capital of the firm and h is the human capital of the worker.

Firms choose their capital stock to maximize profits, before knowing which type of worker will apply to their job. For simplicity, I assume that firms do not bear the cost of capital if they decides not to produce with the worker who has applied to the job. I also denote the cost of capital by c.

Their expected profits are therefore given by

$$\phi x^{H} (1-\beta) \left(k^{1-\alpha}\eta - ck\right) + (1-\phi) x^{L} (1-\beta) \left(k^{1-\alpha} - ck\right),$$

where x^{j} is the probability, chosen by the firm, that it will produce with a worker of type j conditional on matching that type of worker. Therefore, the first term is profits conditional on matching with a skilled worker, and the second term gives the profits from matching with an unskilled worker.

There can be to different types of equilibria in this economy:

1. A pooling equilibrium in which firms choose a level of capital and use it both of skilled and unskilled workers. We will see that in the pooling equilibrium inequality is limited.

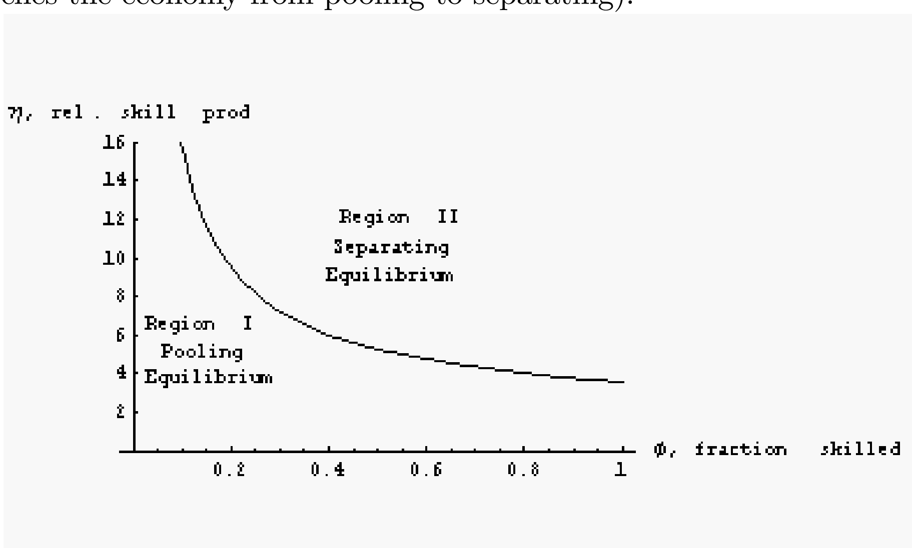
2. A separating equilibrium in which firms target the skilled and choose a higher level of capital. In this equilibrium inequality will be greater.

In this one-period economy, firms never specifically target the unskilled, but that outcome arises in the dynamic version of this economy.

Now it is straightforward to characterize the firms profit maximizing capital choice and the resulting organization of production (whether firms will employ both skilled and unskilled workers). It turns out that first choose the pooling strategy as long as

$$\eta < \left(\frac{1-\phi}{\phi^{\alpha}-\phi}\right)^{1/\alpha}$$

Therefore, a sufficiently large increase in η (in the relative productivity of skilled workers) and/or in ϕ (the fraction of skilled workers in the labor force) switches the economy from pooling to separating).



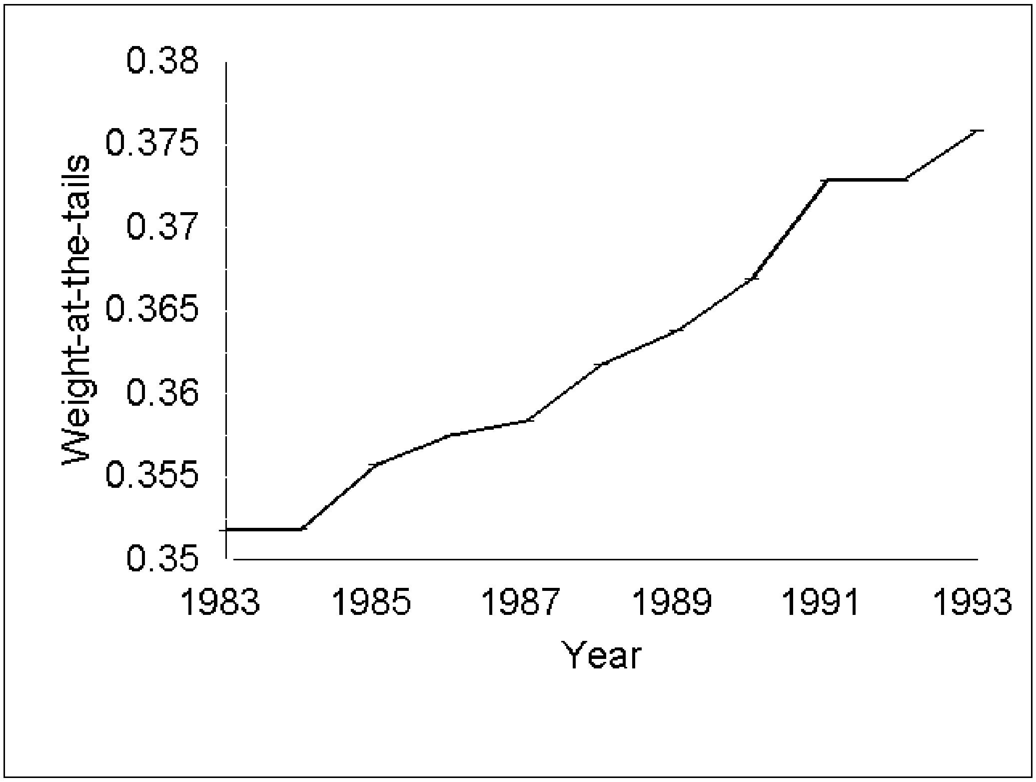
Such a switch will be associated with important changes in the organization of production, an increase in inequality, and a decline in the wages of low-skill workers.

Is there any evidence that there has been such a change in the organization of production? This is difficult to ascertain, but some evidence suggests that there may have been some important changes in how jobs are designed and organized now.

First, firms spend much more on recruiting, screening, and are now much less happy to hire low-skill workers for jobs that they can fill with high skill workers.

Second, as already mentioned above, the distribution of capital to labor across industries has become much more unequal over the past 25 years. This is consistent with a change in the organization of production where rather than choosing the same (or a similar) level of capital with both skilled and unskilled workers, now some firms target the skilled workers with high-capital jobs, while other firms go after unskilled workers with jobs with lower capital intensity.

Third, evidence from the CPS suggests that the distribution of jobs has changed significantly since the early 1980s, with job categories that used to pay "average wages" have declined in importance, and more jobs at the bottom and top of the wage distribution. In particular, if we classify industry-occupation cells into high-wage the middle-wage and low-wage ones (based either on wages or residual wages), there are many fewer workers employed in the middle-wage cells today as compared to the early 1980s, or the weight-at-the-tales of the vob quality distribution has increased substantially as the next figure shows.



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).

This framework also suggests that there should be better "matching" between firms and workers now, since firms are targeting high skilled workers. Therefore, measures of mismatch should have declined over the past 25 or so years. Consistent with this prediction, evidence from the PSID suggests that there is much less over- or under-education today than in the 1970s.

13.3.3 Other approaches the organizational change

Other possible organizational stories could also account for the simultaneous changes in the organization of production and inequality. One possibility is that the introduction of computers has enabled firms to reorganize production, giving much more power to skilled workers, and therefore increasing their productivity. Caroli and Van Reenen provide evidence consistent with this view from British and French establishments. They show that measures of organizational change, such as decentralization of authority, have a strong predictive power for the demand for skills at the establishment level, even after controlling for other determinants of the demand for skills, such as computers.

Another possibility, suggested by Thesmar and Thoenig, is that firms have been gradually changing their organization from mechanistic organizations, which are highly productive at a given task, but not adaptable to changing environments, towards more adaptive organizations, which may be less efficient at a given task, but can quickly and adapt to changes. They link this switch in organizational form to globalization and to the increased availability of skilled workers (as in the above story). Because adaptive organizations require more skilled workers, this change in organizational form increases the demand for skills.

14 Changes in 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 chapter that an explanation for this pattern requires models with multi-dimensional skills.

14.1 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 above analysis showed that there was little evidence in favor of such sizable composition effects. I therefore conclude that the single index model cannot explain the changes in residual inequality during the 1970s and 1980s.

14.2 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.

14.3 Churning and Residual Inequality

Another approach emphasizes that workers of all levels of education may face difficulty adapting to changes. 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. Recent paper by Aghion, Howitt and Violante, for example, suggest that only some workers will be able to adapt to the introduction of new technology, and this will increase wage inequality.

Here I present a simple aversion to give the basic idea. Suppose that the production function of the economy is

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

where H denotes skilled workers, who have productivity A_h, and L denotes unskilled workers who have productivity A_l. As in our baseline model, the skill premium, ω, is given by the standard equation as above, which I repeat here:

$$\omega = \frac{w_H}{w_L} = \left(\frac{A_h}{A_l}\right)^{\rho}\left(\frac{H}{L}\right)^{-(1-\rho)} = \left(\frac{A_h}{A_l}\right)^{(\sigma-1)/\sigma}\left(\frac{H}{L}\right)^{-1/\sigma}. \tag{83}$$

Note that as before, the skill premium is decreasing in H/L.

Workers are not permanently skilled or unskilled, but instead switch between being skilled and unskilled stochastically. For concreteness, suppose that a worker becomes skilled at the flow rate λ (in continuous time), and loses his skills at the flow rate μ. Then in steady state, we have

$$L = \frac{\mu}{\mu + \lambda} \text{ and } H = \frac{\lambda}{\mu + \lambda}.$$

Now consider a change in technology such that both A_l and A_h increase proportionately, but because this technology is different from existing technologies, μ—the rate at which skilled workers lose their skills—increases. This change in adaptability will increase L and reduce H. As a result, the skill premium ω, as given by equation (83), will also increase.

Therefore, in this theory it is the temporary increase in "churning" or dislocation that is responsible for the increase in inequality. This approach also predicts that as workers adapt to the new technology, inequality should fall. Although there is some evidence that residual inequality is no longer growing (see the figure above), there is as yet no evidence of a fall in inequality, despite the very large increase in the supply of skills.

An advantage of this approach is that it is in line with the increased earnings instability pointed out by Gottschalk and Moffit. However, as pointed out a number of times above, 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. 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.

14.4 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 models, somewhat reminiscent of the view of human capital consisting of a number of different attributes as in the Gardener view. In addition, these different dimensions of skills have to correspond loosely to observed and unobserved skills, and be imperfect substitutes (see Acemoglu, 1998). In particular, suppose that there are four types of workers, differentiated by both 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

$$\frac{w_{Ls}}{w_{Lu}} = \left(\frac{A_{ls}}{A_{lu}}\right)^{\rho} \phi_l^{-(1-\rho)} \text{ and } \frac{w_{Hs}}{w_{Hu}} = \left(\frac{A_{hs}}{A_{hu}}\right)^{\rho} \phi_h^{-(1-\rho)}. \tag{84}$$

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. 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 (84) 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.

15 Cross-Country Inequality Trends

As noted above, inequality increased much less in continental Europe than in the U.S. and other Anglo-Saxon economies. There is currently no consensus for why this has been so. There are a number of candidate explanations. These include:

1. The relative supply of skills increased faster in Europe.

2. European wage-setting institutions prevented wage inequality from increasing.

3. For exogenous or endogenous reasons, technical change has been less skill biased in Europe.

The first two are the traditional explanations, put forth by a number of authors. Although they are plausible, there is relatively little work that shows that these explanations can account for the differential inequality trends. I will suggest some preliminary ways of looking into this issue.

15.1 Traditional Explanations

15.1.1 Basic Ideas

The first explanation claims that the more rapid increase in the relative supply of skills in Europe accounts for the lack of increase in inequality there. The second explanation, on the other hand, emphasizes the role of European wage-setting institutions. According to this explanation, it is not the differential growth of skilled workers in the population, but the differential behavior of skilled employment that is responsible for differences in inequality trends across countries. More specifically, firms respond to wage compression by reducing their demand for unskilled workers, and the employment of skilled workers (relative to that of unskilled workers) increases in Europe compared to the U.S. As a result, the market equilibriates with a lower employment of unskilled workers compensating for their relatively higher wages in Europe.

The figure illustrates the first explanation using a standard relative-supply-demand diagram, with relative supply on the horizontal axis and the relative wage of skilled workers, ω, on the vertical axis. For simplicity, I drew the relative supply of skills as vertical. The diagram shows that an increase in the demand for skills, for a given supply of skills, will lead to higher wage inequality. At a simple level, we can think of this economy as corresponding to the U.S., where the consensus is that because of skill-biased technical change or increased trade with less skill-abundant countries, the relative demand for skills grew faster than the relative supply during the recent decades. As a

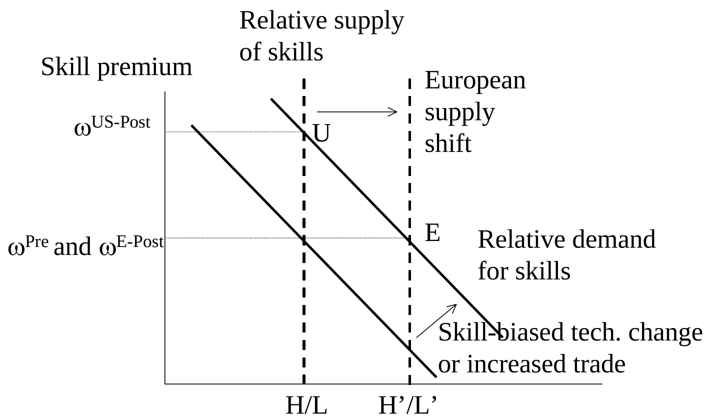


Figure 2:

result of the increase in the relative demand for skills, the skill premium rises from ω^{Pre} to $\omega^{\text{US-Post}}$.

Now imagine that continental Europe is also affected by the same relative demand shifts, but the relative supply of skills also increases. This captures the essence of the first explanation, where the supply of skills increases faster in continental Europe than in the U.S. Then the "European" equilibrium will be at a point like E which may not exhibit greater inequality than before. In fact, the next figure depicts the case in which there is no change in the skill premium in Europe.

Probably the more popular explanation among economists and commentators is the second one above (e.g., Krugman, 1994, OECD, 1994, Blau and

Kahn, 1996). To capture this story, imagine that wage-setting institutions in Europe prevent wage inequality from increasing—for example, because of union bargaining, unemployment benefits, or minimum wages that keep the earnings of low-skill workers in line with those of high-skill workers. This can be represented as an institutional wage-setting line different from the relative supply curve as drawn in the next figure. (To make the story stark, I drew the institutional wage-setting line as horizontal). The equilibrium now has to be along this institutional wage-setting line, and consquently *off the relative supply curve*, causing unemployment. Now, even in the absence of an increase in the relative supply of skills, the skill premium may not increase; instead there will be equilibrium unemployment. In the figure, relative unemployment caused by the increase in the demand for skills is shown as the gap between the relative supply of skills and the intersection between relative demand and institutional wage-setting line. Notice that in the simplest version of the story, there is full employment of skilled workers, and the indicated gap simply reflects unskilled unemployment. The fact that unemployment increased in Europe relative to the U.S. is often interpreted as evidence in favor of this explanation.

But in this context note that in contrast to the prediction of this simple story, unemployment in Europe increased for all groups, not simply for the low-education workers. See, for example, Nickell and Bell (1994), Card, Kramartz and Lemieux (1996) and Krueger and Pischke (1998). Nevertheless, some of the increase in unemployment among the high-education workers in

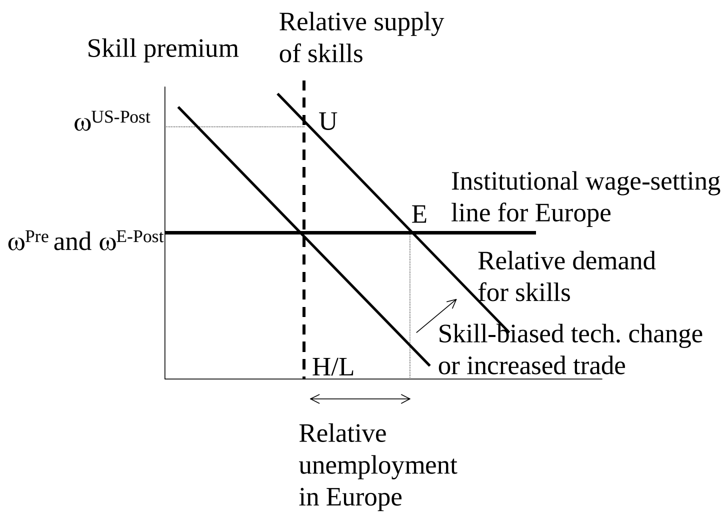


Figure 3:

Europe may reflect the effect of wage compression within education groups on job creation (e.g., if firms are forced to pay the same wages to low-skill college graduates as the high-skill college graduates, they may stop hiring the low-skill college graduates, increasing unemployment among college graduates).

Notice that both of these explanations are "supply-side". Firms are along their relative demand curves, and different supply behavior or institutional characteristics of the European economies pick different points along the relative demand curves. This observation gives us a way of empirically investigating whether these explanations could account for the differential inequality trends, while assuming that relative demand shifts have been similar across countries—that is, that all OECD countries have access to more or less a common set of technologies.

15.1.2 Empirical investigation

Suppose that the aggregate production function for economy j is

$$Y^j(t) = \left[\left(A_l^j(t)\, L^j(t)\right)^\rho + \left(A_h^j(t)\, H^j(t)\right)^\rho\right]^{1/\rho}, \tag{85}$$

which is a multi-country generalization of our above framework.

Suppose that wages are related linearly to marginal product: $w_H^j(t) = \beta MP_H^j(t)$ and $w_L^j(t) = \beta MP_L^j(t)$. The case where $\beta = 1$ corresponds to workers being paid their full marginal product, with no rent sharing. Irrespective of the value of β, we have

$$\omega^j(t) \equiv \frac{w_H^j(t)}{w_L^j(t)} = \frac{MP_H^j(t)}{MP_L^j(t)}.$$

That is, in this specification firms will be along their relative demand curves. Then, we can write

$$\ln \omega^j(t) = \frac{\sigma - 1}{\sigma} \ln\left(\frac{A_h^j(t)}{A_l^j(t)}\right) - \frac{1}{\sigma} \ln\left(\frac{H^j(t)}{L^j(t)}\right). \tag{86}$$

This equation shows that the skill premium is decreasing in the relative supply of skilled workers, $H^j(t)/L^j(t)$, except in the special case where $\sigma \to \infty$ (where skilled and unskilled workers are perfect substitutes).

Let us start with a relatively weak form of the common technology assumption. In particular, suppose that

$$A_h^j(t) = \eta_h^j \theta^j(t)\, A_h(t) \text{ and } A_l^j(t) = \eta_l^j \theta^j(t)\, A_l(t). \tag{87}$$

This assumption can be interpreted as follows. There is a world technology represented by $A_h(t)$ and $A_l(t)$, which potentially becomes more or less skill-biased over time. Countries may differ in their ability to use the world technology efficiently, and this is captured by the term $\theta^j(t)$. Although the ability to use world technology is time varying, it is symmetric between the two sectors. In addition, countries may have different comparative advantages in the two sectors as captured by the terms η_h^j and η_l^j (though these are assumed to be time invariant).

Substituting (87) into (10), we obtain

$$\ln \omega^j(t) = c^j + \ln a(t) - \frac{1}{\sigma} \ln \left(\frac{H^j(t)}{L^j(t)} \right), \tag{88}$$

where $\ln a(t) \equiv \frac{\sigma - 1}{\sigma} \ln \left(A_h(t) / A_l(t) \right)$ is the measure of skill-biased technical change, and $c^j \equiv \frac{\sigma-1}{\sigma} \eta_h^j / \eta_l^j$.

Then, using U.S. data we can construct an estimate for the change in $\ln a(t)$, denoted by $\Delta \ln \hat{a}(t)$, using an estimate for the elasticity of substitution, σ as:

$$\Delta \ln \hat{a}(t) = \Delta \ln \omega^0(t) + \frac{1}{\sigma} \Delta \ln \left(\frac{H^0(t)}{L^0(t)} \right),$$

where $j = 0$ refers to the U.S.

Now define Δ_k as the k-period difference operator, i.e.,

$$\Delta_k x \equiv x(t) - x(t-k).$$

Then, predicted changes in the skill premium for country j between between

$t-k$ and t are given by:

$$\Delta_k \ln \hat{\omega}^j(t) = \Delta_k \ln \hat{a}(t) - \frac{1}{\sigma}\Delta_k \ln\left(\frac{H^j(t)}{L^j(t)}\right). \tag{89}$$

The implicit assumption in this exercise is that there is no delay in the adoption of new technologies across countries. Instead, it is quite possible that some of the new skill-biased technologies developed or adopted in the U.S. are only introduced in continental Europe with a lag. That is, instead of (87), we would have

$$A_h^j(t) = \eta_h^j \theta^j(t) A_h\left(t-k^j\right) \text{ and } A_l^j(t) = \eta_l^j \theta^j(t) A_l\left(t-k^j\right), \tag{90}$$

implying that there is a delay of k^j periods for country j in the adoption of frontier technologies.

Motivated by the possibility of such delays, as an alternative method I use U.S. data from 1974 to 1997 to recover estimates of $\Delta \ln \hat{a}(t)$, and calculate the average annual growth rate of $\ln \hat{a}(t)$, denoted by $\tilde{g}$. I then construct an alternate estimate for the predicted change in the skill premium in country j between dates $t-k$ and t as:

$$\Delta_k \ln \tilde{\omega}^j(t) = \tilde{g}k - \frac{1}{\sigma}\Delta_k \ln\left(\frac{H^j(t)}{L^j(t)}\right) \tag{91}$$

In this exercise, I use 1974 as the starting point, since it is five years prior to the earliest observation for any other country from the LIS data, and five years appears as a reasonable time lag for diffusion of technologies among the OECD countries. I use 1997 as the final year, since this is the final year for which there is LIS data for a country in my sample.

Whether the relative-supply-demand framework provides a satisfactory explanation for cross-country inequality trends can then be investigated by comparing the predicted skill premium changes, the $\Delta_k \ln \hat{\omega}^j(t)$'s from (89) and the $\Delta_k \ln \tilde{\omega}^j(t)$'s from (91), to the actual changes, the $\Delta_k \ln \omega^j(t)$'s.

An empirical investigation along these lines reveals that the changes in relative employment of skilled and unskilled workers go a long way towards explaining the differential developments in inequality and returns to skills across different countries. Nevertheless, it appears that the demand for skills increased less in continental Europe than in the Anglo-Saxon economies. This motivates the theoretical investigation of the reasons for why the demand for skills may change differentially indifferent economies.

15.2 Differential Changes in the Relative Demand for Skills

15.2.1 Basic idea

An alternative to the traditional explanations involves differential changes in the relative demand for skills across countries. These differential changes could reflect four distinct forces:

1. Different countries could develop their own technologies, with different degrees of skill bias.

2. Some countries could be lagging behind the world technology frontier, and may not have adopted the most recent skill-biased technologies.

3. While all countries face the same technology frontier, some may have adopted more skill-biased technologies from this frontier.

4. Different countries have experienced different degrees of trade opening, affecting the demand for skills differentially.

We have already seen that increased international trade is probably not the major cause of the increase in inequality. This leaves us with the first three options. Plausibly, many advanced economies develop some of their own technologies. Nevertheless, it appears plausible that most OECD economies have access, and even relatively rapid access, to the same set of technologies. This suggests that the most likely reason why the relative demand for skills may have behaved differently in continental Europe is not differential development of new technologies or slow technology diffusion, but different incentives to adopt available technologies.

Let me here briefly summarize how the interaction of technology adoption and differences in labor market institutions may induce differential skill-biased technical change in different countries. I will provide a full model of this once we see the general equilibrium search and matching models later in the class.

The basic idea of the theory I propose is to link the incentives to adopt new technologies to the degree of compression in the wage structure, which is in part determined by labor market institutions. In particular, institutional wage compression in Europe makes firms more willing to adopt technologies complementary to unskilled workers, inducing less skill-biased technical change there. This theory is based on three premises:

1. There is some degree of rent-sharing between firms and workers, for example, because of bargaining over quasi-rents.

2. The skill bias of technologies is determined by firms' technology choices.

3. A variety of labor market institutions tend to increase the wages of low-skill workers in Europe, especially relative to the wages of comparable workers in the U.S.

The new implication of combining these three premises is that firms in Europe may find it more profitable to adopt new technologies with unskilled workers than their American counterparts. This is because with wage compression, firms are forced to pay higher wages to unskilled workers than they would otherwise do (that is, greater than the "bargained" wage). This creates an additional incentive for these firms to increase the productivity of unskilled workers: they are already paying high wages, and additional investments will not necessarily translate into higher wages. Put differently, the labor market institutions that push the wages of these workers up make their employers *the residual claimant* of the increase in productivity due to technology adoption, encouraging the adoption of technologies complementary to unskilled workers in Europe.

A simple numerical example illustrates this point more clearly. Suppose that a worker's productivity is 10 without technology adoption, and 20 when the new technology is adopted. Assume also that wages are equal to half of the worker's productivity, and technology adoption costs 6 (incurred solely by the firm). Now without technology adoption, the firm's profits are equal to $1/2 \times 10 = 5$, while with technology adoption, they are $1/2 \times 20 - 6 = 4$. The firm, therefore, prefers not to adopt the new technology because of the subsequent rent-sharing. Next suppose that a minimum wage legislation requires the worker to be paid at least 9. This implies that the worker will be paid 9 unless his productivity is above 18. The firm's profits without technology now change to $10 - 9 = 1$, since it has to pay 9 to the worker because of the minimum wage. In contrast, its profits with technology adoption are still 4. Therefore, the firm canal prefers to adopt the new technology. The reason for this change is clear: because of the minimum wage laws, the firm was already forced to pay high wages to the worker, even when his marginal product was low, so it became the effective residual claimant of the increase in productivity due to technology adoption.

This reasoning implies that there may be greater incentives to invest in technologies complementing workers whose wages are being pushed up by labor market institutions. Since European labor market institutions increase to pay of low-skill workers, technology may be endogenously less skill biased in Europe than in the U.S.

15.3 A simple formalization

Let me now present a simple formalization. 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 a 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.