CENTER FOR LABOR ECONOMICS
UNIVERSITY OF CALIFORNIA, BERKELEY
WORKING PAPER NO. 7

Just Can't Get Enough: More On Skill-Biassed Change and Labour Market Performance

Marco Manacorda* and Alan Manning**

*University College London and Centre for Economic Performance, London School of Economics
**London School of Economics and Centre for Economic Performance, London School of Economics

Revised June 1998

Abstract

It is common to hear the argument that poor labour market performance in OECD countries in recent years is the result of shifts in relative demand against less-skilled workers. But, there is much dispute about whether these trends have been occurring and, if they have, how important they are in quantitative terms. In part these problems come from the absence of a clear conceptual framework in which to think about these issues. In this paper we propose such a framework, propose a measure of skill mismatch that is independent of the definitions of skill and demonstrate using data from a number of countries how it can be used to assess the importance in skill-biassed change in understanding labour market changes in recent years. Our findings suggest that while increased skill mismatch does seem to have occurred in the US and UK, it has not occurred in the other European countries in our sample.

Keywords: Unemployment, Wage Inequality
JEL Classification: J31, J64

We would like to thank Phillip Evans for the most valuable of suggestions, John Schmitt for help with the US data and John van Reenen for helpful comments.

Introduction

It is common in some circles to argue that the problems of labour market performance in the OECD countries have been caused primarily by shifts in relative demand against the less-skilled (see for example OECD, 1994, or Krugman, 1994). Yet others have questioned whether this is really a plausible explanation (see for example, Card, Kramarz and Lemieux, 1994, Nickell and Bell, 1994, or Jackman et al, 1996). These studies typically examine how the relative fortunes of high and low skill individuals have evolved over time and compare these relative fortunes across countries. Education is probably the most commonly used measure of skill but others have been used e.g. the wage in Card, Kramarz and Lemieux (1994).

Although the basic data used in these papers is much the same, the way in which it is used and the conclusions drawn are often quite different. For example, Krugman (1994) points to the larger rise in unemployment rates for the unskilled in Europe, while Nickell and Bell (1994) point out that relative unemployment rates by skill show similar trends across countries and the OECD (1994) look at trends in relative unemployment rates by quartile in the skill distribution. The existing literature does not provide very clear guidance on what evidence should be given most weight and it is difficult to know quite what to make of these disparate pieces of evidence. One of the contributions of this paper is to provide a conceptual framework which can be used to provide guidance as to the way in which one should analyse the basic data. It is doubtful if this paper will be the final word on the subject but it does have the virtue of making clear the assumptions on which the analysis is based.

One of the more serious concerns about the existing papers in this area is that a lot rests on the assumption that levels of education are comparable both across countries and over time. There is good reason to be sceptical about this. While the OECD has put considerable effort into standardizing measures of educational attainment (the ISCED definitions) the fact that these definitions are continually being revised is an indication of the difficulty if not the impossibility of the task. The International Adult Literacy Survey (OECD, 1996) suggests large differences in literacy levels across countries even among individuals with the same ISCED level of education. In addition, the increase in educational attainment in most countries means that a high-school drop-out today is likely to be at a rather different position in the ability distribution today compared to 25 years ago. In this paper we try to achieve comparability across countries and over time by focussing, not on the fortunes of individuals with a given level of education, but on the fortunes of those at a given position in the skills distribution. This has the advantage that it is natural to compare an individual at a given percentile in the skills distribution in one country with someone at the same position in another. Of course, published statistics do not come in this form and one of the contributions of the paper is to show how one can (with certain assumptions) use data on labour market performance by levels of education to make inferences about the performance at different points in the skills distribution.

The plan of the paper is as follows. In the next section we present the basic framework and show how it is the gap between the demand for skills and the supply of skills that is important for labour market performance and proposes a one-dimensional measure of skill mismatch. The second section then shows that, under reasonable assumptions, an increase in this measure of skill mismatch would be expected to be associated with some combination of rising wage inequality and higher unemployment. The third section then shows how this index of skill mismatch can be estimated and the fourth and fifth sections contain empirical evidence on the extent of skill mismatch in six countries. We conclude that there is no evidence of an increase in skill mismatch in the continental European countries but there is evidence of an increase in the US and UK. The sixth section then shows that the measured increase in skill mismatch in the US and the UK is of a magnitude sufficient to explain the rise in wage inequality in these countries.

1. The Theoretical Model: Labour Demand

Let us assume that there is a single index of skill (or human capital) denoted by h (an approach that is also taken by Card and Lemieux, 1996, though their use of the assumption is very different from ours). Suppose the production function is given by the following version of a Cobb-Douglas (we will discuss briefly below the consequences of having a more general specification of the production function):

$$Y(t) = A(t).\exp\left(\int \alpha(h,t)\log N(h,t)dh\right) \tag{1}$$

where N(h,t) is the employment at date t of those with human capital h. This should be thought of as a long-run 'reduced form' production function after one has concentrated out the profit-maximising choice of other inputs so it makes sense to assume that there are constant returns in labour. Assuming that the labour market is perfectly competitive, (1) then leads to the following familiar labour demand curve:

$$\begin{aligned} \log W(h,t) &= \log\alpha(h,t) + \log Y(t) - \log N(h,t) \\ &= \log\alpha(h,t) + \log W(t) + \log N(t) - \log N(h,t) \end{aligned} \tag{2}$$

where variables without h arguments denote aggregate variables. The restrictions on $\alpha(h,t)$ are that the integral with respect to h should sum to one so that density functions are a useful source of possible functions for $\alpha(h,t)$. To keep matters simple (although much of what follows can be done more generally) let us suppose that $\alpha(h,t)$ has the following form:

$$\alpha(h,t) = \phi(h-\mu_{\alpha t}) \tag{3}$$

where $\phi(.)$ is the standard normal distribution. The assumption that the 'variance' of this distribution is one is simply a normalisation which scales the units of h. This specification implies that at each moment in time there is a 'most desired' level of skill which possibly changes over time. It implies, for example, that the demand for brain surgeons is extremely low when development is low as no-one has the other requisite technology to allow them to do their job but that the demand will rise through time as technology advances and then, if we push further on, will then decay again as their skills become superceded by technology.

(2) is written in terms of a given level of skill, h. But, as we described in the introduction, we are interested in what is happening to someone at a given position in the skills distribution which we will denote by F. To say, anything about this we need to bring in the supply of skills. We will assume that the distribution of h in the population is given by:

$$h_{it} \sim N(\mu_{st},\sigma_h^2) \tag{4}$$

An increase in educational attainment will be represented in this framework by an increase in μ_{st}: we will not attempt to model the supply of skills to the economy though any complete model of the economy obviously should do so. Let us use β(h,t) to denote the density of the population with skills h at time t. We can then write the labour demand curve (2) as:

$$\log W(h,t) = \log W(t) + \log[\alpha(h,t)/\beta(h,t)] + \log[(1-u(t))] - \log[1-u(h,t)] \tag{5}$$

where u(h,t) is the unemployment rate of those with skills h and u(t) is the aggregate unemployment rate. But (5) is written in terms of a given level of skill and we want to write this equation in terms of a given position in the skills distribution. Given (4), we have that:

$$h(F,t) = \mu_{st} + \sigma_h \Phi^{-1}(F) \tag{6}$$

which can be used, together with (5) and (3) to derive:

$$\log W(F,t) - \log W(t) = \log\phi(\mu_{st} - \mu_{\alpha t} + \sigma_h \Phi^{-1}(F)) - \log\phi(\Phi^{-1}(F)) + \log\sigma_h - \log(1-u(F,t)) + \log(1-u(t)) \quad (7)$$

where Φ is the distribution function of the standard normal distribution. (7) can be thought of as presenting a trade-off for workers at position F between their log relative wage and their log relative employment rate. The slope of this trade-off is -1 from the Cobb-Douglas assumption and the first-terms on the right-hand side of (7) determine the position of this trade-off so can be thought of as an index of relative demand for the person at position F. Let us denote this relative demand index by log D(F,t). It is worth noting that this demand index must integrate to one so that it makes sense to talk about it as a relative demand index: to see this note that from (5) it is (α/β) which is the relative demand index and that:

$$\int_o^1 \frac{\alpha(h(F,t),t)}{\beta(h(F,t),t)} dF = \int_{-\infty}^{+\infty} \frac{\alpha(h,t)}{\beta(h,t)} \beta(h,t) dh = \int_{-\infty}^{+\infty} \alpha(h,t) dh = 1 \quad (8)$$

where the first equality follows by changing the variable of integration from F to h. How this relative demand index varies with F depends on σ_h and $(\mu_{\alpha t}-\mu_{st})$ which we will denote by μ_t. The case $\sigma_h=1$ is particularly simple and, as we shall see later, also seems to have some empirical relevance[1]. Then the demand index reduces to:

[1] The case where $\sigma_h \neq 1$ is more complicated as then the densities for demand and supply do not have the same variance. In this case the relative demand index will not be monotonic in F which makes things intuitively more difficult and also might cause one to wonder about the suitability of the model.

$$D(F,t) = \frac{\phi(\mu_{st}-\mu_{\alpha t}+\Phi^{-1}(F))}{\phi(\Phi^{-1}(F))} = \frac{\phi(-\mu_t+\Phi^{-1}(F))}{\phi(\Phi^{-1}(F))} = e^{-\frac{1}{2}(\mu_t)^2 + \mu_t\Phi^{-1}(F)} \quad \textbf{(9)}$$

so that relative demand is increasing (decreasing) in human capital as $\mu_t \geq (\leq) 0$ which can be interpreted as the demand for skills running ahead or behind the supply of skills. If we were in the situation where relative demand was decreasing in human capital we would expect there to be low or even negative returns to education so that investment would decrease, reducing μ_{st} and bringing us back towards the case where $\mu_t \geq 0$. So, we might expect that the economically relevant case is where $\mu_t \geq 0$ and our estimates below will suggest very strongly that this is the case.

Notice also that the position of the relative demand curve for a person at a given position in the skills distribution depends on μ_t i.e. on the gap between the demand and the supply for skills. If μ_{st} increases over time (as a result of increasing educational attainment) and $\mu_{\alpha t}$ increases over time (as a result of skill-biassed change) but the gap between them remains the same so that μ_t remains constant then (7) tells us that everyone's relative labour demand curve will remain in the same position and there would be no reason to think that there would be increases in wage inequality and/or unemployment. This makes perfect sense. In spite of the fact that it is a question one often hears raised, it does not make much sense to ask whether there have been any shifts in relative demand against the less-skilled as there probably have been ever since the industrial revolution began (see Goldin and Katz, 1996, for evidence relating to the beginning of this century). More pertinent is the question whether the shifts in relative demand for skilled labour have been matched by equivalent changes in relative supply.

Now consider what is likely to happen if the demand for skills increases faster than the supply so that μ_t increases. By differentiating (7) one can see that such a change will improve the position of the relative demand curve for those at the top of the skills distribution and reduce it for those at the bottom. Again, this makes intuitive sense. This discussion suggests using $\mu_t \equiv (\mu_{\alpha t} - \mu_{st})$ as our index of skill mismatch, an index which is particularly simple. So, if we can figure out a way of estimating μ_t then we can examine the way it changes over time to consider whether there have been any changes in skill mismatch: we show how to do this below.

One way of thinking about (9) is that the relative demand index D(F,t) is a measure of the economic opportunity facing people at a given position in the skills distribution which may be 'taken' in the form of higher relative wages or relative employment or both (see (5)). (7) says that the relative demand index is log-normally distributed over the skills distribution with mean equal to one and variance equal to the square of the mismatch index. So the mismatch index is related to how economic opportunity varies across the skill distribution.

Nothing we have done so far allows us to say what will be the effect of changes in skill mismatch on the aggregate unemployment rate and we cannot say anything about this without discussing the process of wage determination. So, it is to this that we turn next.

2. The Theoretical Model: Wage Determination

In this section we will work with a generic 'wage curve' and we will not go too far into the microfoundations for it although one can justify the type of set-up we are going to use in a bargaining model or an efficiency wage model. The previous discussion showed that if μ_t is constant then the relative labour demand curves do not shift and there is no need for any change in wage inequality or relative employment rates. But, a constant μ_t is only a necessary condition for there to be no change: whether there is any change depends on the process of wage determination. If we restrict our attention to wage curves in which wages depend only on the own unemployment rate then, for no changes in wage inequality and unemployment rates to occur, it is simple to see that the wage curve must be of the following form:

$$\log W(F,t) = \log W(t) - \gamma(u(F,t),F) \tag{10}$$

so that it is relative wages that are related to unemployment. As we do not want to rule out the possibility of neutral change we will assume that the wage curve is of this form. One way of thinking about this assumption is that it implies that the reservation wage of unemployed workers is indexed to average wages e.g. because unemployment benefits are indexed in this way or because wealth effects or support from other family members tie the living standards of the unemployed to the average wage: these seem quite plausible arguments (indeed, Gregg and Manning, 1997, go further and argue that something like this assumption is necessary to explain the economic history of unemployment and the wage structure).

With this wage curve one could, in principle, solve for the unemployment rates and wage structure (we will do something along these lines below). But, here, we are interested in a more limited question: when will increases in skill mismatch lead to increases in aggregate unemployment and/or wage inequality? The following result provides sufficient conditions for this to be the case.

Result 1:

a. If we define $\mu=\mu_\alpha-\mu_s$, then:

$$\frac{\partial u}{\partial \mu} = -\frac{\int \frac{(-\mu+\Phi^{-1}(F))dF}{\psi(u(F),F)}}{1-\frac{1}{1-u}\int \frac{dF}{\psi(u(F),F)}} \quad \textbf{(11)}$$

where:

$$\psi(u,F) \equiv \gamma_u(u,F)+\frac{1}{1-u} \quad \textbf{(12)}$$

b. If $\mu \geq 0$, $u'(F) \leq 0$, $\gamma_u \geq 0$, $\gamma_{uu} \geq 0$, $\gamma_{uF} \leq 0$, then an increase in skill mismatch will lead to a rise in the aggregate unemployment rate and an increase in wage inequality.

Proof: See Appendix A.

The sufficient conditions provided in this result are all reasonable. Our estimates below do suggest that demand does run ahead of supply. The correlation of unemployment rates with education strongly suggests that individuals with more human capital have lower unemployment rates. All the studies of wage curves (notably Blanchflower and Oswald, 1994) suggest that log wages are a convex function of the unemployment rate. Finally, what evidence there is suggests that wages are less responsive to unemployment for those with higher education. So, it does seem reasonable to believe that an increase in skill mismatch could be partially or wholly responsible for the rise in unemployment.

(11) also illustrates the obvious point that the impact of skill mismatch on unemployment depends on how sensitive are relative wages to unemployment. In the extreme case where we have complete wage flexibility $\gamma_u=\infty$ one can then readily check that skill mismatch has no effect on aggregate unemployment though there will be a large impact on wage inequality. On the

other hand if γ_u is close to 0 then the wage distribution will not change much but there will be a large effect on unemployment. This is, of course, often argued to be the difference between the US and European labour markets.

So far, the analysis has been entirely theoretical. So, while it has shown that there are reasons for thinking that an increase in skill mismatch could have caused a deterioration in labour market performance it has not provided any evidence that such an increase in skill mismatch has occurred: this is the subject of the next section.

3. The Data

In this section we investigate the adequacy of this framework and the evidence for the existence and extent of skill mismatch. The theory of the previous sections has all been in terms of human capital and this is a potential source of problems as we do not observe human capital directly. But, we will show below that one can still make progress even if one only has a variable that is correlated with human capital. For this paper we have used education as the appropriate variable as this is what the other papers in the area have most commonly used and it is readily available.

We have data for six countries: France, Germany, Italy, the Netherlands, the UK and the US. This selection was determined by availability of the relevant data more than anything else but they obviously capture an important sub-set of the OECD countries encompassing a wide range of experiences. For each country we have in each year data on employment, unemployment and wages for four education groups. Details of the data sources are provided in the Appendix. While the definitions of educational attainment are meant to be more or less comparable one can see from Figure 1 that the proportions in the different categories vary a great deal across countries: this should make us wary of comparisons of unemployment rates and wages based on allegedly consistent educational definitions[2]. For most of the countries one can see clear evidence of increasing educational attainment. Figure 2 presents the information on the evolution of the wage bill shares showing once again the increase in the shares of the more educated groups. Wage bill shares obviously combine information on labour force shares, relative employment rates and relative wages so Figures 3 and 4 present the data for the evolution of employment rates relative to the average employment rate and wages relative to average wages. The trends are perhaps not particularly easy to see in these Figures so Table 1 presents estimates of the ten-year trend in relative employment and wages. This confirms that most (though not all) of the European countries show evidence of deterioration in the relative employment of the least-skilled whereas it is in relative wages that the trends are most marked in the US and the UK (a trend that has been documented elsewhere e.g. Katz and Murphy, 1992; Schmitt, 1995). It is worth noting that it is possible for changes in relative wages or employment rates to have the same sign for all education groups: this signals the importance of the changing education composition of the labour force.

Now let us consider how we can use this type of information to investigate the problems caused by skill mismatch.

4. Empirical Evidence: Measuring Skill Mismatch

In this section we show how we can estimate our index of skill mismatch, μ_t, to see whether or not there has been an increase in skill mismatch. For this purpose we are going to use

[2] It is worth noting that the OECD quantitative literacy test suggests that allegedly identical levels of education reflect very different levels of literacy in different countries (OECD, 1995).

information on people with different levels of education. We need to have a model of how human capital is related to education. We will assume that human capital is partly determined by schooling, denoted by s (which we will assume to be a continuous variable), but also by 'ability', denoted by ϵ[3] so that schooling is not perfectly correlated with skills. Assume that the human capital of individual i at date t is given by:

$$h_{it} = s_{it} + \epsilon_{it} \qquad \textbf{(13)}$$

Assume that s and ϵ are joint normally distributed with the following distribution:

$$\begin{pmatrix} s_{it} \\ \epsilon_{it} \end{pmatrix} \sim N\left(\begin{bmatrix} \mu_{st} \\ 0 \end{bmatrix}, \begin{pmatrix} \sigma_s^2 & \rho\sigma_s\sigma_\epsilon \\ \rho\sigma_s\sigma_\epsilon & \sigma_\epsilon^2 \end{pmatrix} \right) \qquad \textbf{(14)}$$

The assumption that ability has mean zero is simply a normalization that can be made without loss of generality. This specification allows for schooling to be correlated with ability (as is sometimes claimed). (13) and (14) obviously lead to (4) with $\sigma^2_h = \sigma^2_s + \sigma^2_\epsilon + 2\rho\sigma_s\sigma_\epsilon$.

In this model education is a continuous variable but in our data we only have discrete educational categories. The way we deal with this is to assume that everyone in a particular educational category in our data has a level of schooling between certain limits which remain constant over time. Effectively, we have an ordered probit model for educational attainment.

---

[3] Most studies in this area implicitly assume that schooling and human capital are perfectly correlated. This has the implication that the lowest wage for people with a certain level of schooling will be above the highest level of wages for someone with a lower level of schooling, an assumption that is violated in the data. Introducing the ability component ensures that there is some overlap in the human capital distributions for people with different levels of education as well as allowing us to consider how sensitive are our results to the assumed correlation between schooling and human capital.

Now consider how we can use this information to measure skill mismatch. Suppose that we have divided the population into two education groups, high and low. We will assume that all those classed as high education have education above some level s. We will show how information on the share of the wage bill going to the low education group can be used to infer something about shifts in demand and supply. Working out the relationship between the share of the wage bill going to the low education group and these trends is not entirely straightforward as the trend in demands is defined relative to h which is only imperfectly correlated with s. But the following result shows that, given the assumptions made, there is a simple expression for the share of the wage bill going to those with education less than s (which is something we have data on).

Result 2: The share of the wage bill going to individuals with education less than s at date t, A_{st}, is given by:

$$A_{st} = \Phi\left(\frac{(s-\mu_{st}) + \frac{\rho_{hs}\sigma_s}{\sigma_h}(\mu_{st}-\mu_{at})}{\sigma_s\sqrt{1-\rho_{hs}^2+\frac{\rho_{hs}^2}{\sigma_h^2}}} \right) \qquad \textbf{(15)}$$

where ρ_{hs} is the correlation coefficient between h and s.

Proof: See Appendix A.

It is convenient to invert (15) to give:

$$\frac{(s-\mu_{st})+\frac{\rho_{hs}\sigma_s}{\sigma_h}(\mu_{st}-\mu_{at})}{\sigma_s\sqrt{1-\rho_{hs}^2+\frac{\rho_{hs}^2}{\sigma_h^2}}} = \Phi^{-1}(A_{st}) \tag{16}$$

Note that (16) implies that $\Phi^{-1}(A_{st})$ can be written as an education -specific constant which does not vary over time and a time effect which does not vary across education groups which represents a combination of the demand and supply trends. If one just has information on the evolution in the share of the wage bill of two education groups over time one can use (16) to read off the trends in supplies and demands up to a constant. In this case, there is no test of the adequacy of the model specification. But if one has more than two education groups then one can estimate (16) as all groups should have the same time effects. How well this model fits the data is then a test of the adequacy of the framework.

So, one can use (16) and data on wage bill shares accruing to different education groups over time to estimate some weighted combination of the demand and supply trends. This has a number of useful implications. First, it shows how one can potentially use a variable like schooling that is not perfectly correlated with human capital to investigate the existence of biassed demand shocks. This result can obviously also be applied to any other variable that is correlated, albeit imperfectly, with the underlying measure of skill. For example a number of studies (e.g. Berman, Bound and Griliches, 1994; Machin, 1995) investigate changes in the share of the wage bill going to non-production workers. None of these authors believe that white-collar status is the crucial variable for understanding skill-biassed technical change : it is simply that this is the best data available and it is plausibly correlated with human capital. A result like the one reported above provides some more formal justification for the procedure used in these papers . But the result also suggests some degree of caution in interpreting these results might be warranted. Unless the variable under investigation is perfectly correlated with the human capital variable supply trends as well as demand trends appear in (16) so that the estimated trend will be a linear combination of supply and demand shifts and it is something of an act of faith assuming that increases in the share of the wage bill going to non-production workers is primarily indicative of pure demand shifts. To go to the other extreme, suppose that schooling is uncorrelated with human capital. We will then estimate only the supply shifts from estimating an equation like (16) and investigation of the evolution of wage bill shares can tell us nothing about demand shifts.

But a bit more information can be used to enable us to estimate demand and supply trends. Let us assume that we have data on the fraction of the population below education level s which we will denote by B_{st} which must satisfy the formula:

$$B_{st} = \Phi\left(\frac{s-\mu_{st}}{\sigma_s}\right) \tag{17}$$

which can be inverted to give:

$$\frac{s-\mu_{st}}{\sigma_s} = \Phi^{-1}(B_{st}) \tag{18}$$

Substituting (17) into (16) we have that:

$$\Phi^{-1}(A_{st}) = \frac{1}{\sqrt{1-\rho_{hs}^2+\frac{\rho_{hs}^2}{\sigma_h^2}}}\Phi^{-1}(B_{st}) + \frac{\rho_{hs}(\mu_{st}-\mu_{\alpha t})}{\sqrt{\sigma_h^2(1-\rho_{hs}^2)+\rho_{hs}^2}} \tag{19}$$

so that information on wage bill shares and labour force shares can be used to estimate the evolution of supply relative to demand. Given more than two education groups one can use (19) to estimate the coefficient on $\Phi^{-1}(B_{st})$ and also estimate the trends. This also allows a test of the adequacy of the specification of the model as the fit of this equation should be close to perfect if the model is adequate. A low fit would indicate that it is not just the means of the distribution that is shifting over time and/or that the normal distribution is inappropriate.

(19) has a particularly simple form if $\sigma_h=1$ as it then becomes:

$$\Phi^{-1}(A_{st}) = \Phi^{-1}(B_{st}) + \rho_{hs}(\mu_{st}-\mu_{\alpha t}) = \Phi^{-1}(B_{st}) - \rho_{hs}\mu_t \quad \textbf{(20)}$$

Note that the wage bill shares and labour force shares that we use in these regressions are cumulative ones. As, by definition the cumulative shares of the highest education group must be equal to one there is no useful information here so we only use the cumulative shares for the 3 lowest education groups in each year. These inverse cumulative shares will, by construction, be heteroscedastic and correlated within years. While this does not affect the consistency of OLS estimates, the standard errors must obviously take account of this fact. As we show in Appendix B, we would expect the covariance matrix to be of a particular form and we exploit this in our estimation method, which is feasible GLS.

Table 2a gives the results of the model (19) for our sample countries. Note that the measure of skill mismatch derived from these estimates should be independent of the actual educational categories used so does not rely on educational categories being the same for all countries[4] and should produce estimates that are comparable across countries and over time[5].

---

[4] It does require the assumption that the level of schooling corresponding to a given educational category in a country remains the same over time.

[5] Though, we would expect that the efficiency of the estimate will be affected by the educational categories used.

This is a big advantage of the proposed index and estimation method.

There are a number of aspects of these results worth noting. First, the fit of the equation is close to perfect for all countries implying that the data is broadly consistent with the framework we have described (though this should not be taken to mean that it might not also be consistent with other assumptions about functional forms). Secondly the coefficient on the inverse of the labour force shares is very close to one for the US and Germany and slightly lower for the other European countries. From (19) this implies that $\sigma_h=1$ for the US and that $\sigma_h<1$ for Europe (assuming, as seems reasonable, that $\rho_{hs}>0$). While this could be taken to imply that the distribution of human capital is more compressed in Europe than the US, it would seem that (20) is a reasonable approximation to the data. This is extremely convenient as this case is much easier to understand in intuitive terms and we will use it in what follows. Given this, Table 2b estimates the model imposing this restriction.

All the estimated intercepts in both Tables are positive implying that demand is running ahead of supply. This simply says that the share of the wage bill going to high education groups exceeds their share in the labour force because their wages are higher and their unemployment rates are lower than the average. In terms of trends in skill mismatch, the UK and the US show marked increases, France and Italy show a more modest increase in mismatch in the unrestricted model but none in the restricted model, Germany shows no trend and the Netherlands actually exhibits a decline. Taken together, this suggests that there has been an increase in skill mismatch in the Anglo-Saxon countries but none (or at most a more modest one) in Continental Europe.

Our estimates are based on the assumption that a Cobb-Douglas is an adequate representation of the production function. While there are some estimates suggesting that this is not a bad approximation (see Jackman et al, 1996; Manacorda and Petrongolo, 1996) one might wonder how alternative assumptions would alter our results. The natural generalisation is to a CES technology. This involves giving different weights to relative wages and relative employment rates[6]. As most estimates suggest more substitution between labour of different skills than Cobb-Douglas would imply, we should give greater weight to relative wage changes. For those countries where the increase in mismatch shows up largely through a rise in wage inequality this change would tend to increase measures of skill mismatch. So, moving away from a Cobb-Douglas production function would probably make the performance of the US and the UK look even worse relative to the other European countries.

One might also be interested in how robust are our results to the statistical assumptions made about the covariance structure of the errors. Given that the fit of the model is close to perfect, alternative specifications of the covariance structure will make hardly any difference to the estimated coefficients but they could, conceivably, make a big difference to the standard errors of the estimates. Appendix B describes the way in which we allowed a more general covariance structure: but the bottom line is that it makes no substantive difference to the estimates of mismatch nor to the standard errors.

It is tempting to use estimates like those reported in Table 2b to make comparisons of trends in mismatch over time and within countries. However such comparisons are made difficult by the fact that it is not the mismatch index directly that is being estimated in Table 2 but $\rho_{hs}\mu_t$. Only if ρ_{hs} is constant over time or the same across countries can we use the estimates of Table 2 to infer anything about trends in skill mismatch. But, nothing that we have done so far allows us to estimate this parameter. This is not a problem unique to our approach as most

---

[6] Note that Result 2 does not apply exactly if we have a CES production function, a bad but powerful reason for concentrating attention on a Cobb-Douglas production function.

of the existing literature ignores it by making the convenient but wrong assumption that human capital and schooling are the same thing so that the correlation between them is perfect.

We also might want to quantify the effect of the increase in skill mismatch on wage inequality and unemployment i.e. to ask the question what would wage inequality and unemployment rates have been if we had not observed the increase in skill mismatch. To answer this type of question we also need knowledge of ρ_{hs} so that we can convert our estimates into the parameters we need for (9). So, let us consider how we can estimate the crucial parameter ρ_{hs}.

5. Empirical Evidence: The Correlation Between Human Capital and Schooling

The results of the previous section provide an estimate of $\rho_{hs}\mu$ for each country for each year. But from the point of view of measuring skill mismatch it is really μ in which we are interested. So, it we really want to compare indices of measures of skill mismatch across countries and over time we need to have some estimate of ρ_{hs}. This is the purpose of this section though we should emphasize that there is a certain 'back-of-the-envelope' quality to what follows.

Suppose that wages can be represented by a linear function of human capital plus an error term, ϵ, according to the formula:

$$w = \beta_0 + \beta_1 h + \eta \qquad \textbf{(21)}$$

where η should be thought of as the measurement error in our estimates of wages plus, more generally, wage differentials that exist in reality that are not related to differences in human capital (e.g. compensating wage differentials might fall into this category). Suppose we run a regression of log wages on schooling. We can think of the R^2 from this regression as being the fraction of the variance in wages explained by human capital times the fraction of the variance of human capital explained by schooling. This latter variance is given by ρ_{hs}^2 so that we have:

$$R^2 = \frac{\sigma_w^2 - \sigma_\eta^2}{\sigma_w^2}\rho_{hs}^2 \tag{22}$$

or, re-arranging, we have that:

$$\rho_{hs} = R.\sqrt{\frac{\sigma_w^2}{\sigma_w^2 - \sigma_\eta^2}} \tag{23}$$

So, we can use information on the fraction of the variance of wages explained by education plus the fraction of the variance of wages explained by human capital to estimate the correlation between human capital and education. Table 3 presents estimates of the square root of the correlation coefficient, R, when an earnings function is estimated using only education as controls[7]. We report two estimates, one based on the four educational categories used in the analysis above and another based on as fine a decomposition as is available in our data sources. For the formula above, we would like to have education as a continuous variable so the grouping will mean that we are likely to underestimate R to some degree. The extent of the under-estimate is likely to be quite small given that even four categories seem to be able to do quite well in explaining wages relative to the R^2 for the equation with the larger number of educational categories. Note that for all countries the fraction of the variance in observed wages that can be explained by education seems to be rising modestly through time. There are two possible explanations for this: either the correlation of human capital with schooling is rising through time

[7] Note that we positively do not want other variables in these regressions as our approach has conditioned solely on education and all other variables are in the unexplained part of the skills distribution.

or the fraction of the variance of wages that can be explained by human capital is rising. This latter hypothesis might be plausible for countries like the UK and the US where we know that the wage distribution is widening if we thought that measurement error was approximately constant but is not particularly plausible for the continental European countries. The consequence of assuming that ρ_{hs} is increasing is that the increase in skill mismatch comes to seem even more modest in these countries. This reinforces our earlier conclusion that these countries show no indication of any increase in skill mismatch at all.

On the whole the fraction of the variance in earnings that can be explained by schooling is surprisingly similar in all countries. This would be consistent with the view that ρ_{hs} is similar in all countries and differences in the estimates of Table 2 can be ascribed to differences in skill mismatch suggesting that the problem is worse in the UK and US than in the other European countries.

6. Can Skill Mismatch Be Responsible for the Rise in Wage Inequality?

For countries other than the UK and the US there is obviously little point in assessing formally whether skill mismatch can be held responsible for the rise in unemployment and/or wage inequality for the simple reason that our estimates do not suggest that there has been any rise in skill mismatch. But, for the UK and the US we can try to get some idea as to the size of the problems likely to be caused by the estimated increase in skill mismatch. To do this we also need an estimate of the second term in (23). Let us be conservative and assume that the only variation in wages not explained by human capital is measurement error. Bound and Krueger (1991) provide an estimate that 20% of the observed variance in annual earnings in the CPS is error[8]. If we use this estimate we need to multiply the estimates of R in Table 3 by 1.12 to get an estimate of ρ_{hs}. Note that this estimate is not very sensitive to changes in this proportion. We then get an estimate of the change in μ for the US from -0.68 in 1979 to -0.80 in 1991. Turning to the UK, assuming that the measurement error in wages is of a similar magnitude to that in the US (something on which we have little information) the implied change in μ from 1979 to 1992 is from -0.47 to -0.67 a change of 0.2.

How can we get some idea of the likely impact of these changes? From (7) we can work out how the wage structure should change if relative employment rates were constant i.e. if all the increases in mismatch went into increased wage inequality. If this was the case, we would have:

$$\Delta \log(W(F)/W) = -\frac{1}{2}\Delta\mu^2 + \Delta\mu.\Phi^{-1}(F) \tag{24}$$

We can use this to compare the change in wages at two positions in the wage distribution F_0 and F_1 so that we have:

$$\Delta \log(W(F_1)/W(F_0)) = \Delta\mu[\Phi^{-1}(F_1)-\Phi^{-1}(F_0)] \tag{25}$$

Figure 5 plots the actual change in wages relative to the median against the predicted change assuming that all the rise in skill mismatch went into a rise in wage inequality and relative employment rates were constant. That is, it plots the left-hand side of (25) against the right-hand side where F_0=0.5. The general impression for both the US and the UK is that the actual and predicted changes are quite close though there is some deviation at the extremes of the

[8] It should be remembered that, for most countries, we are using weekly earnings not annual earnings so that there results may not be strictly applicable to our data and we should think of the number used here as a 'best guess'

distribution[9]. This suggests that the increase in skill mismatch is a potential explanation of the rise in wage inequality in both countries.

How should we interpret these changes in the labour markets in US, UK and Continental Europe? The most conventional explanation of the UK and US findings is that there have been shifts in relative demand against the less-skilled. Presumably the same shifts have been occurring in Continental Europe so, if these countries show no evidence of increased skill mismatch it seems likely that the response of the supply of skills has been greater in Continental Europe. However it is worth mentioning that there is another hypothesis that is equally consistent with the data.

Some recent research, notably diNardo, Fortin and Lemieux (1996), has suggested that the declining importance of institutions like unions and the minimum wage played an important part in explaining the rise in wage inequality in the United States. The problem with this conclusion in a conventional competitive model of the labour market is that one would then expect to see a rise in the relative employment of the affected workers with no overall change in skill mismatch: it is very hard to see this in the data. However, if one was to follow the conclusions of researchers like Card and Krueger (1995) into the impact of the minimum wage on employment then one can reconcile the employment and wage movements. Perhaps the simplest way to do this in a framework similar to the one we have used in this paper is to assume that the labour market is monopsonistic and that declines in minimum wages and unions made

[9] Some notes of caution here. First, the observed wage distribution includes the bit due to measurement error that is not taken account of in these computations: this is likely to be small. Second, the predicted wage distribution relates to the distribution among all individuals, the actual data to that among workers in employment. If, as seems likely, employment rates fall as we move up the skills distribution, a given percentile in the employed wage distribution will be a higher percentile in the skills distribution of the population as a whole. However there is no guidance about the direction of the impact of this effect on the change in wages relative to the median.

the labour market more monopsonistic for less-skilled workers. For the usual monopsony reasons such a change would lead to a decline in relative wages and/or relative employment for the affected groups i.e. this hypothesis is observationally indistinguishable from the hypothesis of relative demand shifts[10]. In this situation, the lack of an increase in skill mismatch in the Continental European countries would simply be the result of the fact that they have not had the institutional change in the labour market experienced by the US and the UK.

Our evidence does suggest that the increase in skill mismatch as we have measured it is of a magnitude sufficient to explain the rise in wage inequality in the US and the UK. But, our method cannot identify the reason for this increase in skill mismatch: we have argued that relative demand shifts and institutional change (combined with a non-competitive view of the labour market) are equally consistent with the observed data. And, because we have not found any evidence of increased skill mismatch in many of the European countries, we have not shed any light on the cause of high unemployment rates in these countries beyond suggesting that skill mismatch is not a plausible explanation.

7. Conclusions

In this paper we have proposed a conceptual framework for thinking about the labour market consequences of changes in relative demand and skills of different skills. We have proposed comparing the fortunes of individuals at different positions in the skill distribution arguing that cuts through the problem of assuming comparability of educational classifications across countries and over time. We have shown how one can use data on changes in labour force shares, employment rates, relative wages and the wage structure to provide a measure of the

[10] This argument is presented more formally at the end of Appendix C.

extent of skill mismatch in the economy which is comparable across countries with different measures of educational attainment. Using this technique we found an increase in skill mismatch in the US and the UK but not in the other European countries. So, while the continental countries have had large rises in unemployment, one cannot blame this rise on a change in skill mismatch and one must look elsewhere for explanations. For the UK and the US, the estimated increase in skill mismatch does seem to be of a magnitude capable of explaining the increase in wage inequality.

Appendix A

Proof of Result 1

Combining (10) with (7) we have that:

$$-\gamma(u(F),F) \quad = \quad \ln\phi(-\mu+\Phi^{-1}(F))-\ln(\phi(\Phi^{-1}(F))-\ln(1-u(F))+\ln(1-u) \qquad \textbf{(26)}$$

Let us consider the effect of a change in μ. Differentiating (26) we have that:

$$-\gamma_u(u(F),F)\frac{\partial u(F)}{\partial \mu} \quad = \quad -(-\mu+\Phi^{-1}(F))+\frac{1}{1-u(F)}\frac{\partial u(F)}{\partial \mu}-\frac{1}{1-u}\frac{\partial u}{\partial \mu} \qquad \textbf{(27)}$$

Collecting terms we have that:

$$\frac{\partial u(F)}{\partial \mu} \quad = \quad -\frac{(-\mu+\Phi^{-1}(F))}{\gamma_u(u(F),F)+\dfrac{1}{1-u(F)}}-\left(\frac{1}{1-u}\frac{\partial u}{\partial \mu}\right)\frac{1}{\gamma_u(u(F),F)+\dfrac{1}{1-u(F)}} \qquad \textbf{(28)}$$

Let us define:

$$\psi(u,F) \quad \equiv \quad \gamma_u(u,F)+\frac{1}{1-u} \qquad \textbf{(29)}$$

Now as $u=\int u(F)dF$, we have that:

$$\frac{\partial u}{\partial \mu}\left[1-\frac{1}{1-u}\int\frac{dF}{\psi(u(F),F)}\right] \quad = \quad -\int\frac{(-\mu+\Phi^{-1}(F))dF}{\psi(u(F),F)} \qquad \textbf{(30)}$$

Now as $\psi(u(F),F)>1/(1-u(F))$ the term in square brackets must be positive so the sign of the effect of a change in μ on the unemployment rate depends on the sign of the right-hand side of (30). As $\psi>0$ the term involving μ must be positive if $\mu>0$. Now $\int\Phi^{-1}(F)dF=0$ so we have that:

$$\int\frac{(\Phi^{-1}(F))dF}{\psi(u(F),F)} \quad = \quad Cov\left(\Phi^{-1}(F),\frac{1}{\psi(u(F),F)}\right) \qquad \textbf{(31)}$$

Now $\Phi^{-1}(F)$ is increasing in F and from (29) we have that:

$$\frac{\partial\psi(u(F),F)}{\partial F} \equiv \left(\gamma_{uu}(u,F)+\frac{1}{(1-u))^2}\right)\frac{\partial u(F)}{\partial F} + \gamma_{uF}(u(F),F) \quad \textbf{(32)}$$

which is positive under the assumptions made proving that increases in skill mismatch increase unemployment.

The impact on wage inequality follows straightforwardly. From the wage curve (10) we can see that the unemployment rate and relative wages must go in opposite directions. We know that the increase in skill mismatch must improve the position of the relative labour demand curve for those at the top of the skills distribution and worsen it for those at the bottom. Hence relative wages must rise at the top of the skills distribution and fall at the bottom.

Proof of Result 2

Let us denote by $\alpha_s(s)$ the share of the wage bill going to workers with schooling s. These workers will have different levels of h: we know that the share of the wage bill going to workers with human capital h is given by $\alpha(h)$. Of these workers a fraction $f(s|h)$ have education s where $f(s|h)$ is the density of s conditional on h. So, we must have:

$$\alpha_s(s) = \int \alpha(h)f(s|h)dh \quad \textbf{(33)}$$

From (13) and (14) we have that:

$$s|h \sim N\left(\mu_s+\frac{\rho_{hs}\sigma_s}{\sigma_h}(h-\mu_s),\sigma_s^2(1-\rho_{hs}^2)\right)$$

$$\textit{where } \rho_{hs}\equiv Corr(h,s)\equiv\frac{\sigma_s+\rho\sigma_\epsilon}{\sigma_h} \quad \textbf{(34)}$$

Putting (3) and (34) in (33) we have that:

$$\alpha_s(s) = \frac{1}{2\pi\sigma_s\sqrt{1-\rho_{hs}^2}}\int e^{-0.5(h-\mu_\alpha)^2} e^{-0.5\left(\frac{s-\mu_s-\frac{\rho_{hs}\sigma_s}{\sigma_h}(h-\mu_s)}{\sigma_s\sqrt{1-\rho_{hs}^2}}\right)^2} dh \tag{35}$$

Let us collect the exponential terms and try to find coefficients (δ_0,δ_1,δ_2) so that:

$$(h-\mu_\alpha)^2 + \left(\frac{s-\mu_s-\frac{\rho_{hs}\sigma_s}{\sigma_h}(h-\mu_s)}{\sigma_s\sqrt{1-\rho_{hs}^2}}\right)^2 = (\delta_0 h-\delta_1)^2 + \delta_2 \tag{36}$$

Equating coefficients we have that:

$$\delta_0^2 = 1 + \frac{\rho_{hs}^2}{\sigma_h^2(1-\rho_{hs}^2)} \tag{37}$$

$$\delta_0\delta_1 = \mu_\alpha + \frac{\frac{\rho_{hs}\sigma_s}{\sigma_h}\left(s-\mu_s\left(1-\frac{\rho_{hs}\sigma_s}{\sigma_h}\right)\right)}{\sigma_s^2(1-\rho_{hs}^2)} \tag{38}$$

$$\delta_2 + \delta_1^2 = \mu_\alpha^2 + \frac{\left(s-\mu_s\left(1-\frac{\rho_{hs}\sigma_s}{\sigma_h}\right)\right)^2}{\sigma_s^2(1-\rho_{hs}^2)} \tag{39}$$

Now:

$$
{}_2 = \mu_\alpha^2 + \frac{\left(s-\mu_s\left(1-\frac{\rho_{hs}\sigma_s}{\sigma_h}\right)\right)^2}{\sigma_s^2(1-\rho_{hs}^2)} - \frac{\sigma_h^2(1-\rho_{hs}^2)}{\sigma_h^2(1-\rho_{hs}^2)+\rho_{hs}^2}\left(\mu_\alpha + \frac{\frac{\rho_{hs}\sigma_s}{\sigma_h}\left(s-\mu_s\left(1-\frac{\rho_{hs}\sigma_s}{\sigma_h}\right)\right)}{\sigma_s^2(1-\rho_{hs}^2)}\right)^2
$$

$$
= \frac{\rho_{hs}^2}{\sigma_h^2(1-\rho_{hs}^2)+\rho_{hs}^2}\mu_\alpha^2 + \frac{\sigma_h^2}{\sigma_s^2(\sigma_h^2(1-\rho_{hs}^2)+\rho_{hs}^2)}\left(s-\mu_s\left(1-\frac{\rho_{hs}\sigma_s}{\sigma_h}\right)\right)^2 +
$$

$$
- 2\frac{\sigma_h\rho_{hs}}{\sigma_s(\sigma_h^2(1-\rho_{hs}^2)+\rho_{hs}^2)}\left(s-\mu_s\left(1-\frac{\rho_{hs}\sigma_s}{\sigma_h}\right)\right)\mu_\alpha =
$$

$$
= \left(\frac{\rho_{hs}}{\sqrt{\sigma_h^2(1-\rho_{hs}^2)+\rho_{hs}^2}}\mu_\alpha - \frac{\sigma_h}{\sigma_s\sqrt{(\sigma_h^2(1-\rho_{hs}^2)+\rho_{hs}^2)}}\left(s-\mu_s\left(1-\frac{\rho_{hs}\sigma_s}{\sigma_h}\right)\right)\right)^2 = \qquad \textbf{(40)}
$$

$$
= \left(\frac{s-\mu_s+\frac{\rho_{hs}\sigma_s}{\sigma_h}(\mu_s-\mu_\alpha)}{\sigma_s\sqrt{1-\rho_{hs}^2+\frac{\rho_{hs}^2}{\sigma_h^2}}}\right)^2
$$

Putting these equations into (35), we have that:

$$
\alpha_s(s) = \frac{1}{2\pi\sigma_s\sqrt{1-\rho_{hs}^2}\cdot\delta_0}e^{-0.5\delta_2}\int\delta_0 e^{-0.5(\delta_0 h-\delta_1)^2}dh =
$$

$$
= \frac{1}{\sqrt{2\pi}\sigma_s\sqrt{1-\rho_{hs}^2}\cdot\delta_0}e^{-0.5\delta_2} = \frac{1}{\sqrt{2\pi}\sigma_s\sqrt{1-\rho_{hs}^2+\frac{\rho_{hs}^2}{\sigma_h^2}}}e^{-0.5\left(\frac{s-\mu_s+\frac{\rho_{hs}\sigma_s}{\sigma_h}(\mu_s-\mu_\alpha)}{\sigma_s\sqrt{1-\rho_{hs}^2+\frac{\rho_{hs}^2}{\sigma_h^2}}}\right)^2} \qquad \textbf{(41)}
$$

where the second line follows from the integral of a standard normal random variable. The final expression is the expression for the density function of a normal random variable given in (15).

Appendix B: The GLS Procedure

In the estimation of a regression like (19) or (20) there are good reasons to believe that the individual observations will be both heteroscedastic and correlated across observations within individual years. We deal with this problem by assuming that, within time periods, the covariance matrix of $[\Phi^{-1}(A)-\gamma\Phi^{-1}(B)]$ is of the form $\sigma^2\Omega$ where γ is the slope coefficient in equation (19) and Ω is derived below. We assume that the covariances between periods are all zero. Our estimation method is to use feasible GLS where our estimate of γ in Table 2a is first derived from OLS. To derive Ω we obviously need the covariance matrices of A and B and the covariance between them.

The Covariance Matrix for B

Let d_{ij} be an indicator variable which takes value one if the educational level of individual i is less than j. Then:

$$d_{ij} = \begin{cases} 1 & \textit{with probability } B_j \\ 0 & \textit{with probability } 1-B_j \end{cases} \tag{42}$$

so that:

$$E[d_{ij}]=B_j \tag{43}$$

$$var[d_{ij}]=E[d_{ij}^2]-E[d_{ij}]^2=E[d_{ij}]-B_j^2=B_j(1-B_j) \tag{44}$$

$$cov[d_{ij}d_{ik}]=E[d_{ij}d_{ik}]-E[d_{ij}]E[d_{ik}]=E[d_{ij}]-E[d_{ij}]E[d_{ik}]=B_j(1-B_k) \qquad j<k \tag{45}$$

Using hats on variables to denote sample estimates, this implies that:

$$E(\hat{B}_j)=E(\textstyle\sum_i \frac{d_{ij}}{L})=B_j \tag{46}$$

where L is labour force. Under the assumption of independence of the d_{ij}'s across individuals, it follows that:

$$var(\hat{B}_j)=var(\textstyle\sum_i \frac{d_{ij}}{L})=\sum_i (var(\frac{d_{ij}}{L}))=\frac{B_j(1-B_j)}{L} \tag{47}$$

$$cov(\hat{B}_j,\hat{B}_k)=cov(\sum_i \frac{d_{ij}}{L},\sum_s \frac{d_{sk}}{L})=\frac{1}{L^2}\sum_i \sum_s cov(d_{ij},d_{sk})=$$
$$=\frac{1}{L^2}\sum_i cov(d_{ij},d_{ik})=\frac{B_j(1-B_k)}{L} \qquad j<k \tag{48}$$

Variance of A

Let α_{ij} be the share of the total wage bill of those with education less than j that is earned by individual i. We have that:

$$\alpha_{ij}=\begin{cases}\frac{w_j(F_i)}{wN} & \textit{with probability } (1-u_j)B_j \\ 0 & \textit{with probability } 1-(1-u_j)B_j\end{cases} \tag{49}$$

where F_i is the position of individual i in the earnings distribution of those with education less than j and $w_j(F_i)$ is the wage associated with position i among those with education less than j. Note α_{ij} will be zero if individual i has education more than j. Then:

$$E[\alpha_{ij}|d_{ij}=1]=(1-u_j)\int \frac{w_j(F_i)}{wN}dF_i=\frac{w_jN_j}{wN}\frac{1}{L_j}=\frac{A_j}{L_j} \tag{50}$$

$$var[\alpha_{ij}|d_{ij}=1]=E[\alpha_{ij}^2|d_{ij}=1]-E[\alpha_{ij}|d_{ij}=1]^2=$$
$$(1-u_j)\int(\frac{w_j(F_i)}{wN})^2 dF_i-(\frac{A_j}{L_j})^2=\frac{(1-u_j)}{(wN)^2}var(w_j)+(\frac{A_j}{L_j})^2(\frac{1-u_j}{u_j}) \tag{51}$$

Now our estimate of A_j can be written as:

$$\hat{A}_j=\sum_i \alpha_{ij}d_{ij} \tag{52}$$

Using the previous results we have that:

$$E(\hat{A}_j)=E(\sum_i \alpha_{ij}d_{ij})=\sum_i E(\alpha_{ij}d_{ij})=\sum_i E(\alpha_{ij}|d_{ij}=1)Pr(d_{ij}=1)=\sum_i \frac{A_j}{L_j}B_j=A_j \tag{53}$$

$$
\begin{aligned}
var[\hat{A}_j] &= var(\textstyle\sum_i d_{ij}\alpha_{ij}) = \sum_i var(\alpha_{ij}d_{ij}) = \sum_i E(\alpha_{ij}^2 d_{ij}^2) - \sum_i [E(\alpha_{ij}d_{ij})]^2 = \\
&= L\,E(\alpha_{ij}^2 d_{ij}) - L\,[E(\alpha_{ij}d_{ij})]^2 = L\,E(\alpha_{ij}^2 | d_{ij}=1)Pr(d_{ij}=1) - L[E(\alpha_{ij}|d_{ij}=1)Pr(d_{ij}=1)]^2 = \\
&\frac{L(1-u_j)B_j}{(wN)^2}var(w_j) + \frac{A_j^2}{L_j(1-u_j)} - \frac{A_j^2}{L} = \frac{L_j(1-u_j)}{(wN)^2}var(w_j) + A_j^2\left(\frac{1}{N_j} - \frac{1}{L}\right)
\end{aligned}
\tag{54}
$$

$$
\begin{aligned}
cov[\hat{A}_j, \hat{A}_k] &= cov(\textstyle\sum_i d_{ij}\alpha_{ij}, \sum_s d_{sk}\alpha_{sk}) = \sum_i cov(\alpha_{ij}d_{ij}, \alpha_{ik}d_{ik}) = \\
&= \textstyle\sum_i E(\alpha_{ij}d_{ij}\alpha_{ik}d_{ik}) - \sum_i [E(\alpha_{ij}d_{ij})][E(\alpha_{ik}d_{ik})] = L\,E(\alpha_{ij}^2 d_{ij}) - L\dfrac{A_j}{L}\dfrac{A_k}{L} = \\
&= \frac{L(1-u_j)B_j}{(wN)^2}var(w_j) + \frac{A_j^2}{L_j(1-u_j)} - \frac{A_jA_k}{L} = \frac{L_j(1-u_j)}{(wN)^2}var(w_j) + A_j\left(\frac{A_j}{N_j} - \frac{A_k}{L}\right)
\end{aligned}
\tag{55}
$$

The Covariance between A and B

$$
cov(\hat{A}_j, \hat{B}_j) = cov(\textstyle\sum_i \alpha_{ij}d_{ij}, \sum_i \dfrac{d_{ij}}{L}) = L\,E(\alpha_{ij}\dfrac{d_{ij}}{L}) - L\,E(\alpha_{ij}d_{ij})E(\dfrac{d_{ij}}{L}) = \dfrac{A_j}{L} - \dfrac{A_j}{L}B_j = \dfrac{A_j}{L}(1-B_j) \tag{56}
$$

$$
\begin{aligned}
cov(\hat{A}_j, \hat{B}_k) &= \textstyle\sum_i cov(\alpha_{ij}d_{ij}, \dfrac{d_{ik}}{L}) = L\,E(\alpha_{ij}d_{ij}\dfrac{d_{ik}}{L}) - L\,E(\alpha_{ij}d_{ij})E(\dfrac{d_{ik}}{L}) = \\
&= E(\alpha_{ij}d_{ij}) - A_jB_k = \frac{A_j}{L}(1-B_k) \qquad j<k
\end{aligned}
\tag{57}
$$

$$
\begin{aligned}
cov(\hat{A}_k, \hat{B}_j) &= \textstyle\sum_i cov(\alpha_{ik}d_{ik}, \dfrac{d_{ij}}{L}) = L\,E(\alpha_{ik}d_{ik}\dfrac{d_{ij}}{L}) - L\,E(\alpha_{ik}d_{ik})E(\dfrac{d_{ij}}{L}) = \\
&= E(\alpha_{ik}d_{ij}) - A_kB_j = E(\alpha_{ij}d_{ij}) - A_kB_j = \frac{A_j}{L} - \frac{A_kB_j}{L} \qquad j<k
\end{aligned}
\tag{58}
$$

We are now in a position to estimate the variance of the estimator.

Generalising the Covariance Matrix

For simplicity, let us consider the restricted regression model, where σ_h=1.

$$\Phi^{-1}(A_{st}) = \Phi^{-1}(B_{st}) + (\rho_{hs}(\mu_{st}-\mu_{\alpha t})) + u_{st} \qquad s=1,\ 2,\ 3 \qquad t=1,\ \ldots,\ T \tag{59}$$

The model can be written in the stacked form over the different educational groups as

$$\Phi^{-1}(\underline{A}_t) = \Phi^{-1}(\underline{B}_t) + \beta_t\,\underline{e} + \underline{u}_t \qquad t=1,\ \ldots,\ T \tag{60}$$

where $\underline{e}$ is a (3,1) vector of ones, $\beta=\rho_{hs}\mu$ and

$$\underline{u}_t = [u_{1t},\ u_{2t},\ u_{3t}]' \tag{61}$$

We assume that the random disturbance term u_t is normally distributed with zero mean, zero autocorrelation but with some possible cross-correlation whose structure is not (necessarily) constant over time. Let us denote the covariance matrix of $\underline{u}_t$ by $\sigma^2\Omega_t$ where Ω_t is the matrix derived earlier in this appendix. Now, it is well-known that we can find a P_t such that

$$\Omega_t = P_t^{-1}(P_t^{-1})' \tag{62}$$

so that we can transform the model to:

$$P_t\Phi^{-1}(\underline{A}_t) = P_t\Phi^{-1}(\underline{B}_t) + P_t\underline{e}\beta_t + \underline{v}_t \tag{63}$$

where:

$$\underline{v}_t = P_t\underline{u}_t \tag{64}$$

Under our assumptions the error term has a diagonal variance covariance matrix

$$var(\underline{v}_t) = \sigma^2 P_t\Omega_t P_t' = \sigma^2 I_3 \tag{65}$$

Our GLS procedure essentially estimates (61). Now consider the alternative, unrestricted, hypothesis that:

$$var(\underline{v}_t) = C \qquad \textbf{(66)}$$

where C is some matrix (not necessarily an identity matrix). The log likelihood of the unrestricted model can be written

$$\ln L(\underline{\beta}, C) = -\frac{3T}{2}\ln(2\pi) - \frac{T}{2}\ln|C| - \sum_t \underline{v}_t' C^{-1} \underline{v}_t \qquad \textbf{(67)}$$

while for the restricted model

$$\ln L(\underline{\beta}, \sigma^2) = -\frac{3T}{2}\ln(2\pi + \sigma^2) - \frac{1}{2\sigma^2}\sum_t \underline{v}_t' \underline{v}_t \qquad \textbf{(68)}$$

A log likelihood ratio test of our assumption about the covariance matrix of $\underline{v}_t$ against the alternative of (66) is then of the form:

$$L = -2\ln\frac{L^R}{L^U} = 3T\ln\hat{\sigma}_v^{R2} - T\ln|\hat{C}| + \frac{1}{\hat{\sigma}^{R2}}\sum_t \hat{\underline{v}}_t^{R\prime}\hat{\underline{v}}_t^R - \sum_t \hat{\underline{v}}_t^{U\prime}\hat{C}^{-1}\hat{\underline{v}}_t^U \qquad \textbf{(69)}$$

where U and R denote respectively maximum likelihood estimates of the unrestricted and the restricted model. Under H_0, the log-likelihood ratio is distributed as a chi-square with a number of degrees of freedom equal to the number of restrictions. A completely unrestricted specification of C cannot be estimated, but we can estimate generalisations of our basic model and test our basic model against them. In doing this we are interested not just in whether we can reject our restricted model but whether the estimates from the unrestricted model have different estimates of the extent of mismatch and different standard errors. In all the specifications we tried, no substantive difference in the results were found.

Appendix C: Monopsony and Skill Mismatch

A crude way of representing a monopsonistic labour market is to note that the first-order condition for profit maximisation can be written as MRPL=W(1+θ) where θ is related to the elasticity of the labour supply curve facing the firm. The 'labour demand curve' of (2) then becomes:

$$\log W(h,t) = \log\frac{\alpha(h,t)}{1+\theta(h,t)} + \log Y(t) - \log N(h,t)$$

$$= \log\frac{\alpha(h,t)}{1+\theta(h,t)} - \log\int\frac{\alpha(h,t)dh}{1+\theta(h,t)} + \log W(t) + \log N(t) - \log N(h,t) \qquad \textbf{(70)}$$

By inspection of (70) we can see that changes in α(h,t) are observationally equivalent to changes in θ(h,t) so that an increase in the relative monopsony power of employers in the low-skill segments of the labour market (e.g. because of a decline in the minimum wage) will be indistinguishable from skill-biassed technical change.

Data Appendix

For each country we divide the labour force into four education groups. For each of these groups we collected data on the share of the labour force in each group, the unemployment rate and wages. Sources and definitions are listed for each country.

France: Data come from *Enquete Emploi*, 1982-1993.The four groups are: up to primary school (*cep, be, beps*), junior high school (*cap, bep*), high school (*bac* or equivalent), university education.

Germany: Data come from the German Socio Economic Panel, 1984-1994. The four groups are: without vocational qualification, with vocational qualification, with apprenticeship, with university education. The wage variable is monthly income form labour. Data are weighted by cross-sectional weights.

Italy: Data on wages come from the Bank of Italy Survey of Household Income and Wealth for the years 1977-1984,1986,1987,1989,1991. Data on employment and unemployment come from the *Annuario statistico italiano*, ISTAT, various issues. The four groups are: up to primary school, junior high school, high school and university degree or above. Wages are defined as take home annual pay. Wage data presented are weighted by the population weights.

Netherlands: Employment data come from *Arbeidskrachtetentelling* for the period 1979-1985 and from *Enquete Beroepssbevolking* for the period 1990-1993. Wage data come from *Tijdreeksee Arbeidsrekeningen*. The four groups are: up to primary school, junior high school, high school and university degree or above. Earning concept: gross monthly wages.

UK: Data come from the General Household Survey for 1974-92. The four groups are: those with no qualifications, those with O-levels or equivalent qualification, those with A-levels or eqivalent qualifiiation and university graduates. The wage variable is weekly earnings.

US: Data come from the Current Population Survey. Data for the period 1973-1978 are derived form the May Annual Files, while data for the period 1979-1991 come from the Outgoing Rotation Group. The four groups are high-school drop-outs, high-school graduates, those with 2 years of college and those with 4 or more years of college. The wage variable is weekly earnings. Since data on earnings are top coded in the CPS we estimate a tobit model of log earings on experience, experience square, four education dummies, a sex dummy, a race dummy, 50 state dummies, a part-time dummy, a dummy for married individuals. Based on the estimated standard deviation of residuals from this tobit regression we then construct an uncensored normal distribution and impute wages for top coded observations on the assumption of a log normal distribution of wages. Wage data are weighted by the earning weights.

Table 1

Trends in Relative Employment and Wages by Education

	Education Category			
	1	2	3	4
Relative Employment Rate				
France	-.024	.010	.006	.022
Germany	-.016	.013	-.009	-.001
Italy	.008	.006	.031	.056
Netherlands	-.017	-.010	-.020	.026
UK	-.023	-.001	.003	.017
US	-.023	-.005	.005	-.001
Relative Wages				
France	-.048	-.027	-.112	-.088
Germany	.007	.001	-.083	-.071
Italy	-.020	-.052	-.055	-.075
Netherlands	.009	.016	-.102	-.119
UK	-.089	-.085	-.104	.013
US	-.113	-.079	-.048	.033

Notes.
The numbers refer to the estimate 10-year trend in the relevant variable. The relative employment rate is the employment rate of the particular education group relative to the aggregate employment rate and the relative wage is the wage of the particular education group relative to the average wage.

Group 1 has the lowest level of educational attainment, group 4 the highest.

Table 2a:GLS Estimates of Shifts in Demand Relative to Supply: Unrestricted Model

country	France	Germany	Italy	Netherlands	UK	US
$\Phi^{-1}(B)$	.894 (.005)	.991 (.008)	.933 (.003)	.970 (.008)	.912 (.011)	1.001 (.004)
1973						.226 (.027)
1974					.191 (.019)	.227 (.027)
1975					.189 (.017)	.259 (.028)
1976					.207 (.017)	.247 (.028)
1977			.050 (.002)		.223 (.016)	.241 (.026)
1978			.072 (.003)		.207 (.017)	.246 (.026)
1979			051 (.003)	.270 (.024)	.210 (.017)	.229 (.014)
1980			.073 (.003)		.230 (.020)	.237 (.013)
1981			.061 (.002)	.262 (.024)	.264 (.019)	.247 (.014)
1982			.066 (.002)		.233 (.021)	.278 (.014)
1983			.064 (.002)	.276 (.025)	.278 (.021)	.286 (.014)
1984	.203 (.002)	.213 (.008)	.084 (.002)		.300 (.023)	.285 (.014)
1985	.198 (.002)	.244 (.010)		.251 (.026)	.308 (.027)	.300 (.014)
1986	.204 (.002)	.230 (.009)	.091 (.002)		.315 (.022)	.309 (.013)
1987	.206 (.002)	.242 (.011)	.093 (.002)		.327 (.022)	.314 (.013)
1988	.212 (.002)	.262 (.011)			.329 (.022)	..311 (.014)
1989	.215 (.002)	.224 (.010)	.089 (.002)		.350 (.022)	.310 (.013)
1990	.227 (.002)	.256 (.010)		.220 (.026)	.356 (.023)	.323 (.013)
1991	.237 (.002)	.238 (.009)	.093 (.002)	.215 (.026)	.359 (.022)	.329 (.013)
1992	.248 (.003)	.232 (.009)		.197 (.026)	.347 (.023)	
1993	.245 (.003)	.212 (.010)		.207 (.027)		

1994	.254 (.003)	.212 (.011)				
R^2	.9999	.9997	.9999	.9988	.9980	.9994
No. Of Observations	33	33	36	24	57	57

Table 2b: GLS Estimates of Shifts in Demand Relative to Supply: Restricted Model

country	France	Germany	Italy	Netherlands	UK	US
$\Phi^{-1}(B)$	1.000	1.000	1.000	1.000	1.000	1.000
1973						.226 (.026)
1974					.206 (.030)	.228 (.027)
1975					.198 (027)	.260 (.027)
1976					.213 (.027)	.247 (.027)
1977			.059 (.019)		.225 (.026)	.242 (.026)
1978			.076 (.015)		.205 (.027)	.246 (.026)
1979			.051 (.016)	.251 (.033)	.208 (.028)	.230 (.014)
1980			.069 (.012)		.224 (.032)	.238 (.013)
1981			.053 (.010)	.251 (.033)	.256 (.030)	.248 (.013)
1982			.054 (.014)		.222 (.031)	.279 (.014)
1983			.049 (.014)	.264 (.035)	.257 (.033)	.286 (.014)
1984	.211 (.010)	.210 (.009)	.065 (.013)		.273 (.033)	.285 (.014)
1985	.204 (.010)	.241 (.010)		.238 (.035)	.289 (043)	.301 (.013)
1986	.205 (.010)	.223 (.009)	.065 (.011)		.279 (.034)	.309 (.013)
1987	.202 (.011)	.239 (.010)	.062 (.012)		.294 (.035)	.314 (.013)
1988	.205 (.010)	.255 (.010)			.293 (.035)	.311 (.013)
1989	.207 (.011)	.220 (.010)	.052 (.011)		.312 (.035)	.310 (.013)
1990	.212 (.012)	.252 (.009)		.208 (.035)	.317 (.037)	.324 (.013)
1991	.217 (.012)	.234 (.009)	.048 (.011)	.200 (.035)	.319 (.034)	.330 (.013)
1992	.225 (.013)	.228 (.009)		.181 (.035)	.303 (.036)	
1993	.217 (.014)	.208 (.009)		.191 (.036)		

1994	.223 (.013)	.207 (.010)				
R^2	.9988	.9997	.9995	.9975	.9948	.9826
No. Of Observations	33	33	36	24	57	57

Table 3
R from a Regression of Log Wages on Education

country	Germany		Italy		UK		US	
1973							.32	.34
1974					.30	.35	.33	.35
1975					.30	.34	.34	.36
1976					.32	.36	.34	.36
1977			.28	.29	.36	.40	.32	.35
1978			.25	.26	.35	.40	.34	.36
1979			.25	.26	.35	.39	.32	.34
1980			.26	.28	.35	.39	.33	.35
1981			.21	.22	.37	.41	.33	.35
1982			.26	.26	.38	.42	.34	.36
1983			.29	.30	.35	.39	.35	.37
1984	.34	.38	.31	.31	.36	.40	.33	.35
1985	.38	.40			.35	.41	.34	.36
1986	.38	.39	.31	.32	.36	.41	.35	.36
1987	.42	.44	.31	.31	.39	.42	.36	.37
1988	.45	.46			.35	.40	.35	.37
1989	.38	.40	.33	.34	.37	.42	.40	.41
1990	.41	.42			.38	.42	.40	.41
1991	.39	.42	.31	.31	.39	.42	.40	.41
1992	.50	.53			.39	.43		
1993	.52	.56	.32	.33				
1994	.48	.52						

Notes.

1. There are no results for France or the Netherlands because we do not have access to micro data for these countries.

2. These numbers are the square root of the R^2 from a regression of log wages on a set of dummy variables for educational attainment alone.

Figure 1
Labour Force Shares by Education

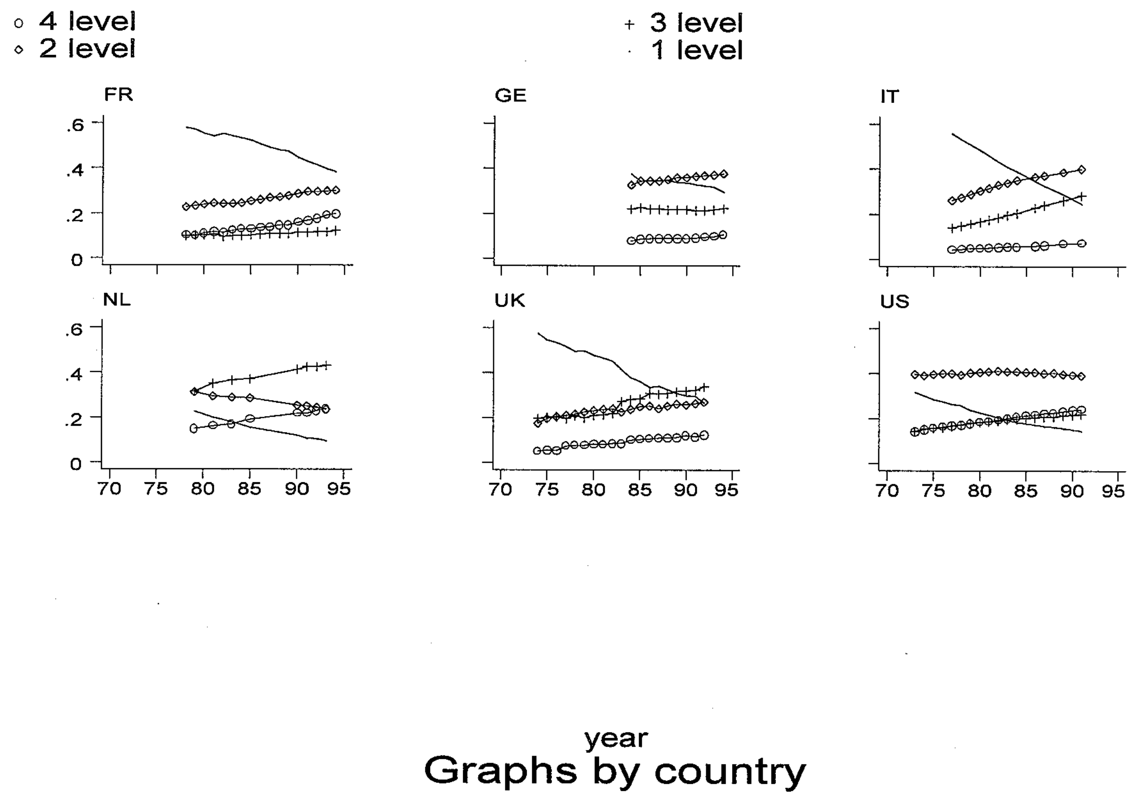


Figure 2
Wage Bill Shares by Education

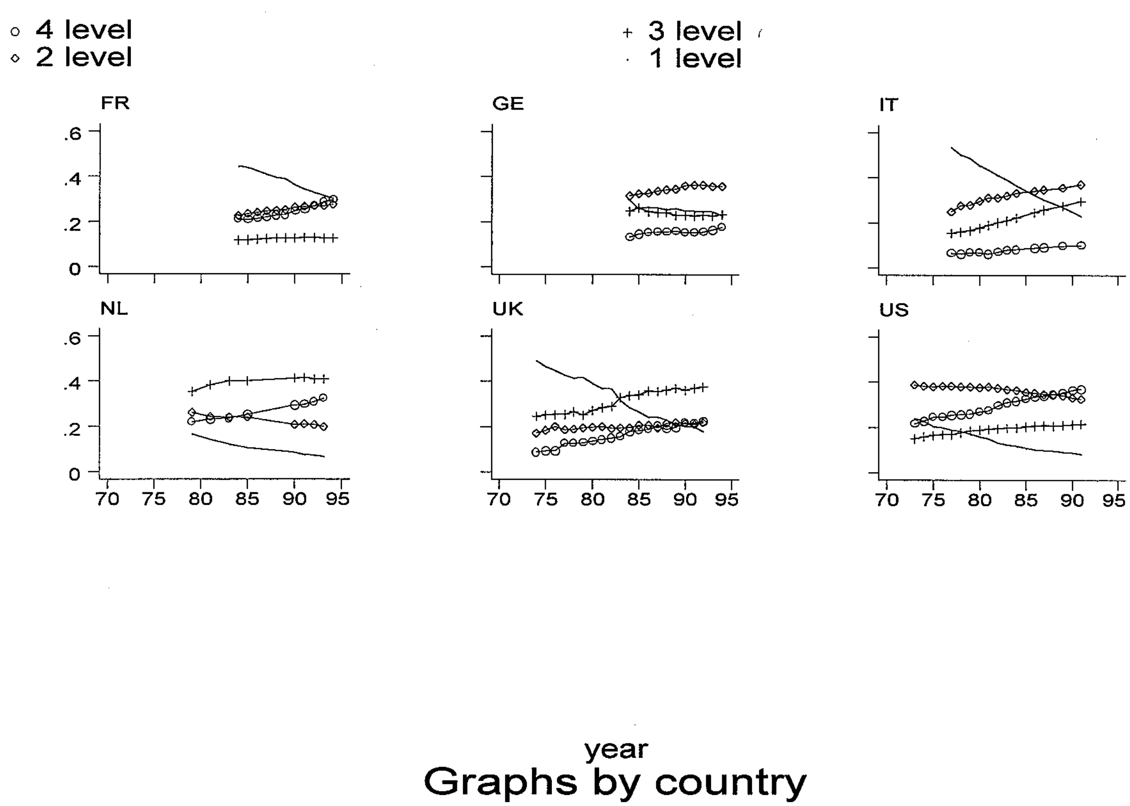


Figure 3
Relative Employment Rates by Education

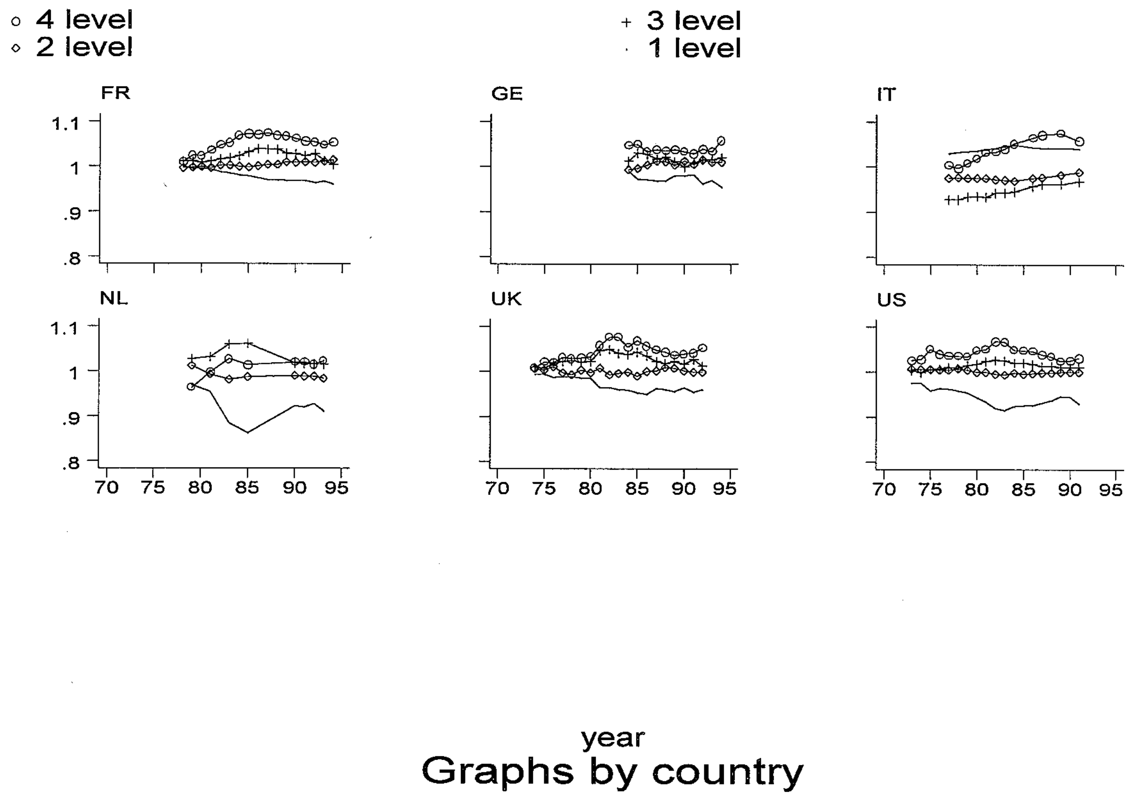


Figure 4
Relative Wages by Education

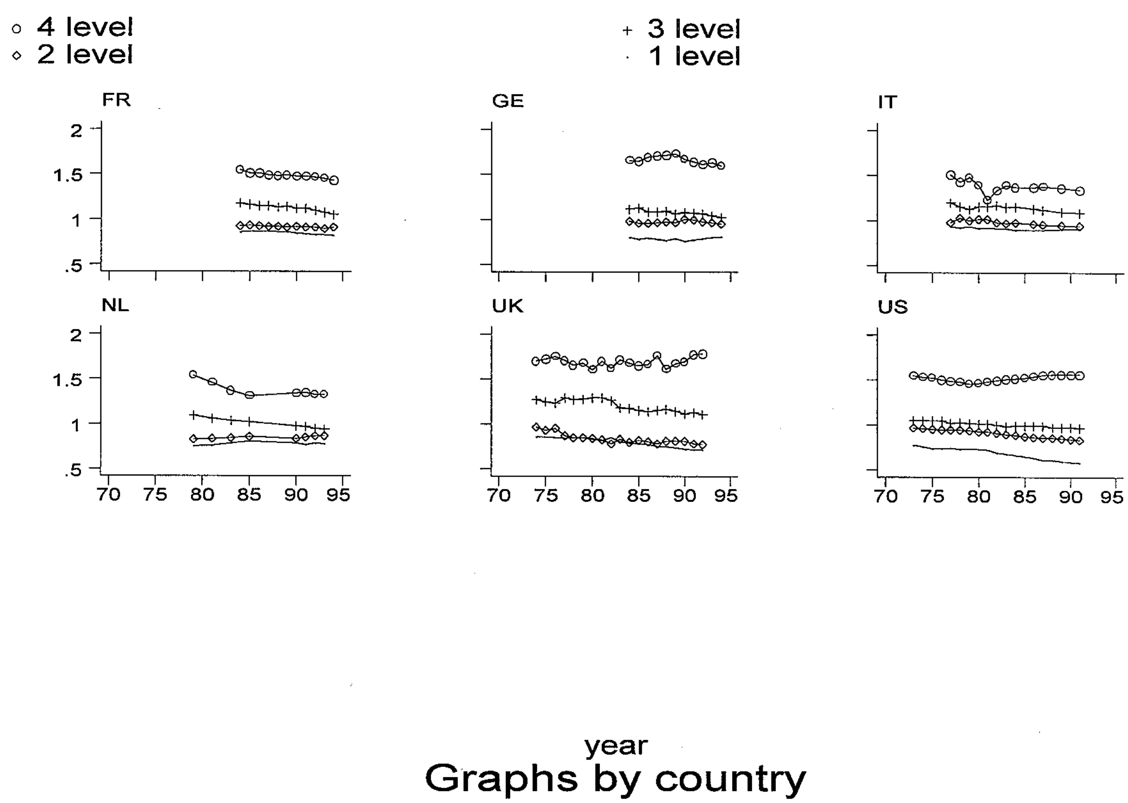
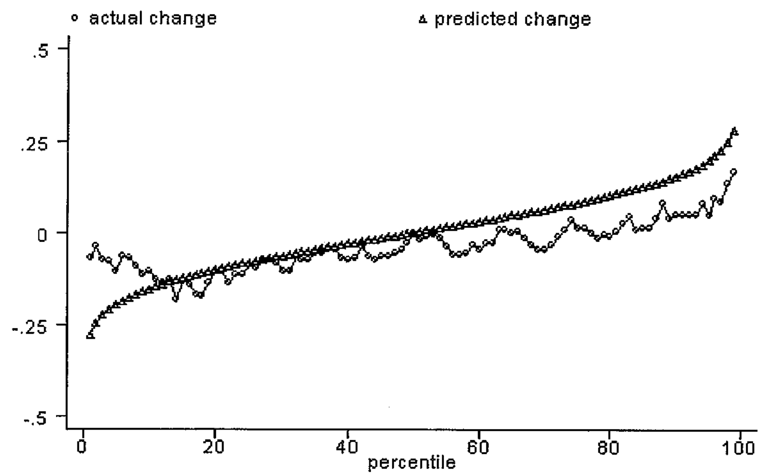
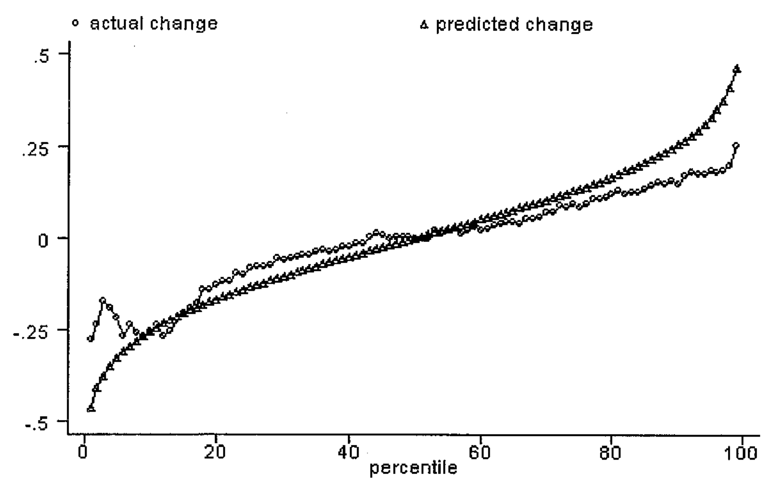


Figure 5
Predicted and Actual Rises in Wage Inequality

United States



United Kingdom



Notes.

1. The actual change is the change in the log wage at different percentiles relative to the log median over the period 1979-91 for the US and 1979-92 for the UK.
2. The predicted change is what would be obtained using the formula in equation (23).

References

Berman, E., J.Bound and Z.Griliches (1994) "Changes in the Demand for Skilled Labor within US Manufacturing Industries: Evidence from the Annual Survey of Manufacturing", Quarterly Journal of Economics, 109, 367-398.

Blanchflower, D. And A.Oswald (1994) The Wage Curve, Cambridge, MIT Press.

Bound, J. and A. Krueger "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?", Journal of Labor Economics, 9, 1-24.

Card, D., F.Kramarz and T.Lemieux (1994) "Changes in the Relative Structure of Wages and Employment: A Comparison of the United States, Canada and France", unpublished, Princeton University.

Card, D. and A.Krueger (1995) Myth and Measurement: The New Economics of the Minimum Wage, Princeton: Princeton University Press.

Card, D. And T.Lemieux (1996) "Wage Dispersion, Returns to Skill, and Black-White Wage Differentials", Journal of Econometrics , 74, 319-61.

DiNardo, J, N.Fortin and T.Lemieux (1996) "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach", Econometrica, 64, 1001-44.

Goldin, C. and L.Katz (1996) "The Origins of Technology-Skill Complementarity", NBER Working Paper No. 5657, July 1996.

Gregg, P. and A. Manning (1997) "Skill-Biassed Change, Unemployment and Wage Inequality", European Economic Review, 41, 1173-120.

Jackman, R., R.Layard, M.Manacorda and B.Petrongolo (1996) "European vs. US Unemployment: Different Responses to the Same Shock?", unpublished, LSE.

Juhn, C., K.Murphy and R.Topel (1991) "Why Has the Natural Rate of Unemployment Increased Over Time? Brooking Papers on Economic Activity, 75-126.

Katz, L. and K.Murphy (1992) "Changes in Relative Wages, 1963-1987: Supply and Demand Factors", Quarterly Journal of Economics, 57, 35-78.

Krugman, P. (1994) "Past and Prospective Causes of High Unemployment", unpublished paper, Stanford University.

Layard, R., S.Nickell and R.Jackman (1991) Unemployment: Macroeconomic Performance and the Labour Market, Oxford University Press: Oxford.

Machin, S. (1995) "Changes in the Relative Demand for Skills in the UK Labour Market", in A.Booth and D.Snower (eds) Acquiring Skills: Market Failures, Their Symptoms and Policy Responses, Cambridge University Press: Cambridge.

Machin, S. (1996) "Wage Inequality in the UK", Oxford Review of Economic Policy, 12, 47-64.

Manacorda, M. and B.Petrongolo (1997) "Skill Mismatch and Unemployment in OECD Countries", forthcoming, Economica.

Nickell, S. and B.Bell (1995) "The Collapse in the Demand for the Unskilled and Unemployment Across the OECD", Oxford Review of Economic Policy, 11, 40-62.

OECD (1994) OECD Jobs Study, OECD:Paris

OECD (1995) Literacy, Economy and Society, OECD: Paris

Schmitt, J. (1995) "The Changing Structure of Male Earnings in Britain, 1974-88" in R.Freeman (ed) Differences and Changes in Wage Structures, University of Chicago Press: Chicago.29