Working Paper Series
No. 3

Measuring skill: a multi-dimensional index

Miguel Portela

September 2000

Published in *Economics Letters*, 2001 (72), 27-32

Núcleo de Investigação em Microeconomia Aplicada
Universidade do Minho

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA E DA TECNOLOGIA

Measuring skill: a multi-dimensional index

Miguel Portela[*][†]

NIMA and University of Minho

Economics Department, Gualtar, 4710-057 Braga, Portugal

4th September 2000

Abstract

Traditionally, skill is measured concentrating on just one dimension of the worker's ability, usually years of schooling or the blue/white collar nature of the job. This paper proposes a measure of skill that combines, in a multiplicative way, several of the observed components of skill, as well as its unobserved dimension. The proposed index is intuitively appealing and it is flexible, in the sense that it can accommodate as many (or as little) dimensions of human capital as feasible and suitable for the analysis to be undertaken.

Keywords: composite index, skill

JEL classification: C23, J71

[*]E-Mail: mangelo@eeg.uminho.pt Tel.: +351-253-604510; fax: +351-253-676375

[†]I am grateful to Ana Rute Cardoso for her encouragement and critical comments. I also wish to thank Anabela Botelho, Lígia Pinto and Paulo Guimarães for their comments.

1 Introduction

The skill of a worker reflects several of the person's attributes, ranging from the formal knowledge acquired in school, to general aptitudes obtained in the market place or specific abilities developed within a firm, together with capacities, innate or developed throughout life, often unobservable. The quantification of the worker's skill and the valuation of its returns have been lying at the core of most of the recent debates in empirical labor economics. A few examples help to clarify the relevance of the topic.

Rising wage inequality from the 1980's onwards has been brought about in several countries by the rising wages of more skilled workers. Kremer and Maskin (1996) analyze the interaction between the skill distribution, the labor force homogeneity within firms and the wage inequality in the economy. They show that the evaluation of the segregation of workers by their skill between firms is relevant to understand the wage inequality in an economy.

Matching theory and economic growth theory also emerge among several examples that could be pointed out to illustrate the relevance of the skill measurement issue. On one hand, the assignment of workers to jobs, or to firms, depends on the worker's skill and on the skill needs of the job. On the other hand, the potential growth of a country depends on the skill of its labor force. According to Mincer (1993, p. 286), "at the macro-economic level the social stock of human capital and its growth are central to the process of economic growth".

Traditionally, labor economics has measured skill using, one of the following alternatives, the education level, the labor market experience, or the distinction between production and non-production workers[1]. Often, the wage itself has been used as an indication of the worker's productivity and therefore his/her productive skills, an idea disputed by non-competitive theories of wage determination.

A measure of skill should, ideally, capture several dimensions of the human capital of the worker. For example, Blau and Kahn (1996) suggest a composite measure of skill using information on schooling and experience. Teulings (1995) also suggests that workers skills should be measured in a continuous one-dimensional scale.

This paper proposes an index of skill that can be useful in a diversity of studies. The proposed index includes different dimensions of a worker's skill, using a specific functional form to reach an index that is intuitively appealing and flexible enough to take into account

[1] See, for example, Bound and Johnson (1992) and Berman *et al.* (1998).

as many (or as little) dimensions of the human capital of the worker as feasible, and suitable for the study to be undertaken.

2 The skill index

A composite index is constructed. The index can synthesize different dimensions of the productivity of the worker, such as schooling, labor market experience and unobservable ability, to cite a few examples.

Note, first of all, that these variables may have different units of measurement and widely different means and degrees of dispersion. Therefore, the variables will not be combined in an additive way into the composite index, but instead in a multiplicative way. The components of skill will be successively introduced into the index, and each new component will complement the information contained in the previous ones, increasing or decreasing their intensity.

Beginning by the observable component, completed years of schooling can be used as a starting point. When entering the labor market, every individual has a certain level of education, which can shape his/her learning ability. Note however that the school level, evaluated as completed years of schooling, is zero for some individuals. For such workers, it is not possible to introduce a multiplicative correction to account for the other dimensions of the skill. Let us then assume, just for the sake of clarifying the explanation, that all workers enter the labor market with the average schooling in the economy, represented below by $mschool$. This average is then corrected to account for the worker's actual schooling, taking into account his/her relative position in the distribution of schooling. For individuals with a schooling level above the average, the starting value must be increased, while those that are below the average will have their initial classification decreased. In a first step, when only the information on schooling is considered, the skill index is the following,

$$S = mschool * a,$$

where a is the correction factor taking account of the actual position of the individual in the schooling distribution. To compute the correction factor a, the cumulative logistic distribution is considered. The choice of this functional form is justified by its properties, namely the marginal impact of the values away from the average of the distribution. For

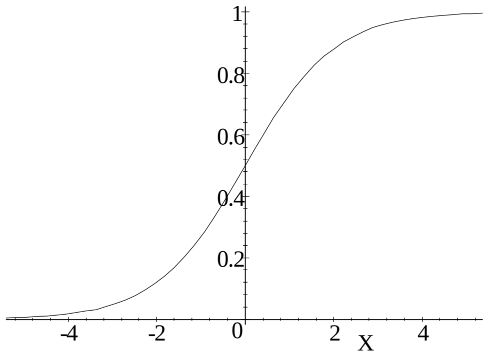

Figure 1: Logistic cumulative distribution function

values close to the average, the function is almost linear in the correction, whereas it moderates the impact of the correction for values away from the average, a topic that will deserve further comment below.

The corrected schooling level for individual i is thus measured as

$$mschool * \left(0.5 + \frac{e^{\frac{school_i - mschool}{sschool}}}{1 + e^{\frac{school_i - mschool}{sschool}}}\right) \tag{1}$$

where $school_i$ stands for the schooling level of worker i, evaluated as years of schooling, $mschool$ represents the average of the school level in the population, and $sschool$ is the standard deviation of schooling. This index represents the base skill of individual i and it is obtained by standardizing the relative position of the worker in the schooling distribution. The correction factor, $0.5 + \frac{e^{\frac{school_i - mschool}{sschool}}}{1+e^{\frac{school_i - mschool}{sschool}}}$, takes values between 0.5 and 1.5. Values above 1 are reached when the schooling level of the worker is above the average in the economy, while levels of schooling below the average are associated with a correction factor of less than 1.

Figure 1 shows the function $\frac{e^x}{1+e^x}$, where, for schooling, $x = \frac{school_i - mschool}{sschool}$. The marginal variations in the correction factor that occur for values of schooling away from the average are reduced. That is to say that schooling levels away from the average differ little between themselves in their effect on the skill of the individual. The idea is that with a low education level the individual's skill is very low, so he gets increasing returns to his investment in schooling. After a while, the returns are still positive - since the logistic distribution is monotonically increasing - but diminishing. What individuals learn in school is mainly calculus and literacy abilities, which means that after a certain point in the schooling distribution what the worker gets in terms of additional productive skills is still positive, but decreasing.

The next step consists in correcting for the experience. Two alternatives of implementation are feasible in this case. Experience can be considered by itself, in which case what is relevant is the deviation of a worker's experience relative to the average experience of the workforce. On the other hand, one can consider the experience as a complement to schooling, comparing only the experience of individuals with the same education level. The choice between these two alternatives depends on the relative importance of each of the two variables considered, schooling and experience. The second alternative is chosen, since it assumes that the impact of experience on the ability of a worker depends on his/her schooling. The same length of experience in the labor market has a different impact on the productive skill, depending on the individual's schooling. After accounting for labor market experience, the skill index becomes

$$S_i = mschool * \left(0.5 + \frac{e^{\frac{school_i - mschool}{sschool}}}{1 + e^{\frac{school_i - mschool}{sschool}}}\right) * \left(0.5 + \frac{e^{\frac{exper_i - mexper_{|school_i}}{sexper_{|school_i}}}}{1 + e^{\frac{exper_i - mexper_{|school_i}}{sexper_{|school_i}}}}\right) \quad (2)$$

where $exper_i$ is the experience of the individual, $mexper_{|school_i}$ is the average experience within schooling level $school_i$, $sexper_{|school_i}$ is its standard deviation , and $\frac{exper_i - mexper_{|school_i}}{sexper_{|school_i}}$ is computed for individual i.

This second correction intends to reflect the fact that individuals with identical schooling levels can have different skills, and that part of this difference can be captured by their experience. Therefore, for individuals with the same schooling level, the skill is higher the greater the positive deviance from the mean experience. However, a higher dispersion of the experience for a group of workers with a certain schooling level, implies that the relevance of this difference will be reduced, since it is a more common situation. The higher the deviance from the mean experience, the lower the impact in the skill that results from variations in experience. The use of the logistic distribution allows to control for the effect of outliers, remaining an almost linear relation in the relevant interval of variation around the mean.

Once again, the value of the correction factor, $0.5 + \frac{e^{\frac{exper_i - mexper_{|school_i}}{sexper_{|school_i}}}}{1+e^{\frac{exper_i - mexper_{|school_i}}{sexper_{|school_i}}}}$, can vary between 0.5 and 1.5. When the experience is below average, the coefficient to multiply for the base skill will be less than 1, representing a *penalty* on the ability of the worker. When the experience is greater than the average, the coefficient will be higher than 1, and it will increase the initial measure of skill associated with the schooling.

The same procedure might be used to introduce other dimensions of the human capital of the worker into the index. Having introduced the effect of the schooling and the individual's experience, we can go further by introducing the knowledge about the worker's non-observed productivity. Where a panel dataset of employees is available, a wage regression with a specific effect for the worker may be run, as presented in the following equation,

$$y_{it} = \alpha_i + \boldsymbol{\beta}'\mathbf{x}_{it} + \varepsilon_{it}, \tag{3}$$

where y_{it} represents the logarithm of the wage for individual i at time t, α_i is the individual effect, which is taken to be constant over time t and specific to individual i, and $\mathbf{x}_{it}$ are K regressors, not including the constant term. Thus, this wage regression isolates the workers' unobserved productivity, which is represented by the individual effect.

Therefore, after correcting the schooling for the experience of the worker, we can proceed using a similar procedure to correct for the dispersion of unobserved productivity inside the same group of schooling and experience. Again, it is implicit that a hierarchy exists between the dimensions of the skill of the individual. The initial dimension is the schooling, followed by the experience, and finally the characteristics that cannot be directly observed. The innate abilities of an individual would have little value for a firm or for an economy if the individual had not been in school. For example, if the worker is not able to read, it would be difficult for him to operate technology. Likewise, experience is important to improve and to give some value to innate abilities. If the worker does not have a training period, his innate abilities would have a small impact on his productivity.

The skill index proposed in this paper is therefore:

$$\begin{aligned} S_i \quad = \quad mschool * \left(0.5 + \frac{e^{\frac{school_i - mschool}{sschool}}}{1 + e^{\frac{school_i - mschool}{sschool}}}\right) * \left(0.5 + \frac{e^{\frac{exper_i - mexper_{|school_i}}{sexper_{|school_i}}}}{1 + e^{\frac{exper_i - mexper_{|school_i}}{sexper_{|school_i}}}}\right) * \\ * \left(0.5 + \frac{e^{\frac{effect_i - meffect_{|school_i, exper_i}}{seffect_{|school_i, exper_i}}}}{1 + e^{\frac{effect_i - meffect_{|school_i, exper_i}}{seffect_{|school_i, exper_i}}}}\right) \end{aligned} \tag{4}$$

where $effect_i$ corresponds to the individual's specific effect estimated in the wage regression specified in (3), $meffect_{|school_i, exper_i}$ it is the average of those effects for the individuals with the same schooling and experience, while $seffect_{|school_i,exper_i}$ is the standard deviation of those effects.

3 Conclusion

This paper proposes a multidimensional measure to evaluate the human capital of the worker. The proposed index presents several advantages relative to the standard measures of skill: it is a standardized measure of skill across variables with different units of measurement; it presents great flexibility to accommodate as many dimensions of the human capital of the worker as adequate for the study at hand; and it has an intuitive interpretation. Several of the topics currently under debate in the labor economics literature can be approached using this index, namely the skill upgrading of a country or of an industry, changes in the composition of labor demand, the segregation of workers between firms and their impact on wage dispersion.

References

Berman, E.; Bound, J.; Machin, S. (1998) "Implications of skill-biased technological change: international evidence", *Quarterly Journal of Economics*, 113(4), 1245-1279.

Blau, F.; Khan, L. (1996) "International differences in male wage inequality: institutions versus market forces", *Journal of Political Economy*, 104(4), 791-837.

Bound, J.; Johnson, G. (1992) "Changes in the structure of wages in the 1980's: an evaluation of alternative explanations", *American Economic Review*, 82(3), 371-392.

Kremer, M.; Maskin, E. (1996) "Wage inequality and segregation by skill", National Bureau of Economic Research Working Paper #5718.

Mincer, J. (1993) *Studies in human capital*, Edward Elgar, Gower House.

Teulings, C. (1995) "The wage distribution in a model of the assignment of skills to jobs", *Journal of Political Economy*, 103(2), 280-315.