Review of Income and Wealth
Series 48, Number 2, June 2002

MEASURING EDUCATIONAL HETEROGENEITY AND LABOR QUALITY: A NOTE

BY MOGENS FOSGERAU

Centre for Economic and Business Research

SVEND E. HOUGAARD JENSEN

*Centre for Economic and Business Research and EPRU**

AND

ANDERS SØRENSEN

Centre for Economic and Business Research

This paper investigates the magnitude of the mismeasurement that occurs when only a few education categories are used in the construction of a constant quality index for labor input. By employing a very comprehensive data set it is found that the error resulting from the omission of information on education is relatively small. The empirical results are thus supportive of the current state of practice of constructing indices of constant quality labor input.

I. INTRODUCTION

Spurred by recent discussions on the role of information technology for economic growth (Jorgenson, 2001), there has been a renewed interest in productivity measurement and growth accounting in particular. Indeed, papers based on the methodology of growth accounting, such as Oliner and Sichel (2000) and Jorgenson and Stiroh (2000), have been widely cited. Also, the recently published manual on productivity measurement by the OECD (2001) marks an important step toward improving the statistical basis for offering better guidance to policy-making. In the same vein, international research networks have been established with the aim of providing comparable growth accounts for a number of countries.[1]

The increase in educational attainment that has taken place in OECD countries and elsewhere puts emphasis on the importance of accounting for the quality of the labor input. Thus, for example, it has been found that increasing labor quality has contributed on average 0.59 percentage points annually to increasing labor input in the U.S. over the period 1948–95, and that nearly all of this trend can be attributed to increasing levels of educational attainment (Ho and Jorgenson, 2001). Accounting for the heterogeneity of labor requires cross-classification

Note: Without implication, we would like to thank an anonymous referee for helpful comments and suggestions. Financial assistance from the Danish Ministry of Economics and Business Affairs and the Danish National Research Foundation is gratefully acknowledged. Opinions expressed are those of the authors, and should not be assumed to be those of our sponsors.

*Corresponding author: CEBR, Langelinie Alle 17, DK-2100 Copenhagen OE, Denmark: E-mail: shj@cebr.dk.

[1]See, for example, http://www.conference-board.org/economics/klems.

of the workforce along a number of dimensions, of which the level of educational attainment is arguably the most important.

The data requirements involved in studies of that kind are severe. Therefore, the decision of how much costs to invest in data acquisition must be balanced against the gains that may be obtained in terms of more precise results. Also, additional data may not be available. Existing studies that we know of have only used a few categories of educational attainment; for example, Ho and Jorgenson (2001) use six different educational groups. In a recent study Timmer (2000) recommends a *minimum* of four categories of educational attainment, thus indicating that more would be desirable. Therefore, the question remains if there are benefits of increasing the number of categories even further. The purpose of this note is to help answer this question.

In pursuing this objective we use data for Denmark of an exceptionally high quality. Indeed, we use a data set with 29 different educational groups, defined by both the level and type of education. Furthermore, the data set encompasses the entire population for the period 1980–98, grouped by nine employment statuses, sex, eight age groups, four job categories and nine groups for working hours. The effect of compositional changes with respect to education is measured after controlling for compositional changes with respect to sex, age and employment class. Two indices are reported. The first index measures the effect of education when only four aggregate levels of education are distinguished. This index is compared to a second index computed using 29 levels and types of education. The results show that only little precision is gained by expanding the number of educational categories. The two indices are very close and highly correlated. Thus, the practice of using only a few categories is supported.

The paper proceeds as follows. The methodology is outlined in Section 2, Section 3 presents results and discussion, and Section 4 concludes. An appendix contains information on our data set and, in order to assess international comparability, a brief description of the education institutional setting and the wage structure in Denmark is presented.

2. Methodology

Our first objective is to establish an index for labor services that takes into account the heterogeneity of labor. Next we compute some decompositions of this index to highlight the role of education in the growth in labor services. We follow the methodology given in, e.g. Jorgenson, Gollop, and Fraumeni (1987).

Labor can be differentiated by a number of socio-economic characteristics as well as by the industry of occupation. Letting B be a set of different labor types, L_b denotes the quantity of labor input of type $b \in B$ and w_b denotes the average rate of labor compensation for labor of type b. We shall define an index of labor services corresponding to the set of labor types in B, treating different types of labor as distinct inputs, i.e.

$$L_B = L_B(\{L_b, w_b\}_{b \in B}). \tag{1}$$

Let $L = \sum L_b$ denote total employment and write the total labor compensation as $wL = \sum_b w_b L_b$. Under cost minimization, labor is compensated according

to its marginal product. Differentiate (1) with respect to time and substitute the first order conditions for cost minimization to obtain an index for the growth in labor services L_B (hatted variables are growth rates):

$$\hat{L}_B = \sum_{b \in B} \frac{w_b L_b}{wL} \hat{L}_b. \tag{2}$$

This index depends on what characteristics are used to distinguish different types of labor. In principle, the set of relevant characteristics is endless. In practice, however, we are restricted. To account for as much heterogeneity as possible we let $B = E \times S \times A \times C$ be the product of sets indexing education, sex, age and employment class and thus obtain our *total index of labor services*, L_{ESAC}. This index adjusts for heterogeneity across all our observed dimensions of heterogeneity including the education dimension.

Dropping the education dimension an index L_{SAC} is obtained, correcting for all dimensions of heterogeneity except education. With this we decompose the total index of labor services, L_{ESAC}, into two elements.

$$\hat{L}_{ESAC} = (\hat{L}_{ESAC} - \hat{L}_{SAC}) + \hat{L}_{SAC} \tag{3}$$

The difference $\hat{L}_{ESAC} - \hat{L}_{SAC}$ is the part of the increase in labor services that results from shifts between different types of educations.

$$\hat{L}_{ESAC} - \hat{L}_{SAC} = \sum_{esac} \frac{w_{esac} - \mathrm{w}_{sac}}{w} \frac{\Delta L_{esac}}{\mathrm{L}}.$$

The effect of education is positive if labor input of some educational type grows (falls) which has wages higher (lower) than the overall average. This is seen by substitution with (2).

The main objective is to analyze the effect of varying the level of detail in the distinction between educations. For that purpose we aggregate the set of educations into four groups, $f \in F$. Define L_{FSAC}, as above. Then

$$\hat{L}_{ESAC} - \hat{L}_{FSAC} = \sum_{f} \sum_{sac} \sum_{e \in E_f} \frac{w_{esac} - w_{fsac}}{w} \frac{\Delta L_{esac}}{L}. \tag{4}$$

Note that this difference is composed of four terms corresponding to the elements of F. Each term is positive if the labor input of some type *esac* grows (falls) and the wage for that type, w_{esac}, is higher (lower) than the average wage for the aggregate education level, w_{fsac}.

3. Results and Discussion

All results pertain to the private sector in Denmark for the period from 1980 to 1998. The data set is described in detail in the data appendix. Here the classification of educations is discussed. Based on the list of educations we have defined our initial grouping of educations from two variables, where one is related to the length or the level of education and the other is related to the type of education. The type of education is a general classification according to subject. We distinguish technical, natural science, social science, humanities, social and health,

trade and office, service, and other. Not all combinations of level with type exist, for example elementary school is classified only as "other." The combination of level with type yields 29 different educations.

The 29 educations are grouped into four aggregate education levels. The grouping is roughly comparable to the U.S. classification into <4 years of high school, completed high school, 1–3 years of college, and $4+$ years of college, which is also the definition of categories recommended by Timmer (2000). The first aggregate group, elementary schooling, comprises 8, 9 and 10 years of elementary school and the group with unknown education; the second aggregate group comprises four types of high school education; the third aggregate group comprises four types of vocational education and five types of short further education; and the fourth aggregate group comprises six types of medium further education and six types of long further education.

Figure 1 shows three indices with 1990 normalized to 1, and the corresponding numbers are given in Table 1. The index of labor services controlling for sex, age and employment class, L_{SAC}, exhibits large cyclical variations over an increasing trend. The cyclical effect is generally due to variations in employment while the trend effect mainly results from the changing age structure of employment.

The point of interest here is to provide a perspective on the two indices of education. The index L_{FSAC}/L_{SAC} is the contribution of education to total labor services measured with four levels of education. Similarly, L_{ESAC}/L_{SAC} is the contribution of education measured with 29 different levels and types of education.

It is evident from Figure 1, that the two indices measuring the contribution of education track each other very closely. The two indices of education grow quite steadily over the period. L_{FSAC}/L_{SAC} grows at an average annual rate of 0.268 percent while L_{ESAC}/L_{SAC} grows at 0.243 percent. The mean growth rate of

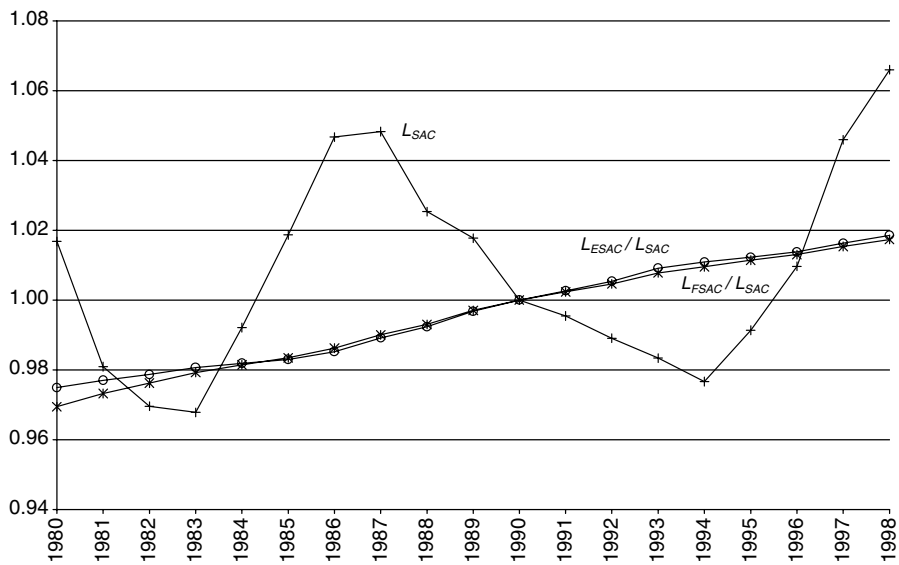

Figure 1. Partial Indices of Labor Services

TABLE 1

PARTIAL INDICES OF LABOR SERVICES

Year	L_{ESAC}/L_{SAC}	L_{FSAC}/L_{SAC}	L_{SAC}
1980	0.9749	0.9695	1.0168
1981	0.9770	0.9732	0.9809
1982	0.9787	0.9762	0.9696
1983	0.9807	0.9792	0.9678
1984	0.9819	0.9814	0.9921
1985	0.9830	0.9835	1.0187
1986	0.9852	0.9862	1.0467
1987	0.9892	0.9901	1.0482
1988	0.9924	0.9930	1.0253
1989	0.9968	0.9970	1.0177
1990	1.0000	1.0000	1.0000
1991	1.0026	1.0023	0.9954
1992	1.0054	1.0046	0.9891
1993	1.0092	1.0078	0.9834
1994	1.0109	1.0096	0.9766
1995	1.0123	1.0114	0.9914
1996	1.0138	1.0130	1.0097
1997	1.0163	1.0154	1.0459
1998	1.0186	1.0174	1.0660

Source: Statistics Denmark (2000) and own calculations.

the difference between the two indices is thus 0.024 percent. In comparison with the contribution from education this is a small difference.

Disaggregating the grouping of educations turns out to decrease the measured contribution of education. As shown in (4) this phenomenon occurs when employment within the four aggregate groups shifts into types with lower than average wages. We shall pursue this point.

The difference between L_{FSAC}/L_{SAC} and L_{ESAC}/L_{SAC} is decomposed into four terms according to the four aggregate education levels, f, in (4). Figure 2 presents the corresponding four indices, and the numbers are given in Table 2. The average annual growth rates over the whole period from 1980 to 1998 are small and not very different. For elementary schooling, high school, and vocational and short further education we find average growth rates of –0.01 percentage points, for medium and long further education we find a small positive average growth rate of 0.003 percentage points. The possibility existed that we would find large positive and negative growth rates for the four indices, netting out to a small aggregate effect. That this turns out not to be the case provides further support for saying that distinguishing between many types of education is less important from a growth accounting perspective.

Three of the four indices exhibit quite steady rates of change, the exception being the index for the group of elementary educations which has decreased substantially up to 1985. Up to 1985, this is due to measurement problems with education not being recorded for a share of people born before 1921. The two indices of education, L_{FSAC}/L_{SAC} and L_{ESAC}/L_{SAC}, grow over the period 1985–98 at average annual rate of 0.261 percent and 0.274 percent, respectively, implying that the difference between the two indices is narrowed compared to the full period. This strengthens the main point of the paper, namely that the standard practice of just using four main types of education is valid.

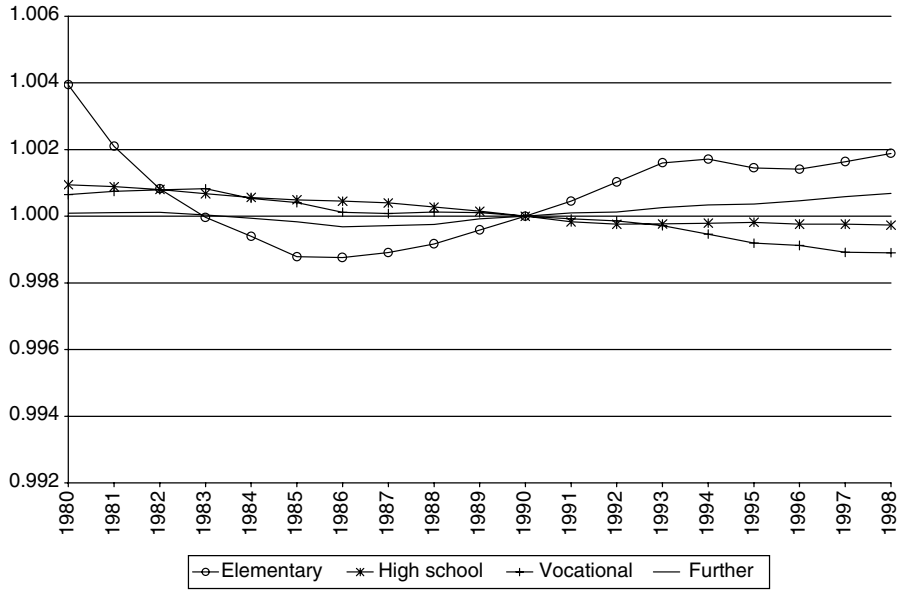

Figure 2. Decomposition of the difference between L_{FSAC}/L_{SAC} and L_{ESAC}/L_{SAC}.

TABLE 2

Decomposition of the Difference Between L_{FSAC}/L_{SAC} and L_{ESAC}/L_{SAC}

Year	Elementary	High School	Vocational	Further
1980	1.0039	1.0009	1.0006	1.0001
1981	1.0021	1.0009	1.0007	1.0001
1982	1.0008	1.0008	1.0008	1.0001
1983	1.0000	1.0007	1.0008	1.0000
1984	0.9994	1.0006	1.0005	0.9999
1985	0.9988	1.0005	1.0004	0.9998
1986	0.9988	1.0005	1.0001	0.9997
1987	0.9989	1.0004	1.0001	0.9997
1988	0.9992	1.0003	1.0001	0.9998
1989	0.9996	1.0001	1.0001	0.9999
1990	1.0000	1.0000	1.0000	1.0000
1991	1.0005	0.9998	0.9999	1.0001
1992	1.0010	0.9998	0.9999	1.0001
1993	1.0016	0.9998	0.9997	1.0003
1994	1.0017	0.9998	0.9995	1.0003
1995	1.0014	0.9998	0.9992	1.0004
1996	1.0014	0.9998	0.9991	1.0005
1997	1.0016	0.9998	0.9989	1.0006
1998	1.0019	0.9997	0.9989	1.0007

Source: Statistics Denmark (2000) and own calculations.

4. Concluding Remarks

Data collection is often both time consuming and expensive. Whether such an effort is worthwhile can be difficult to determine in advance. The results presented in this paper provide some guidance in judging such considerations.

More specifically, we have investigated the effect of increasing the level of detail on education in the construction of a composition-adjusted quantity index of employment. Such indices have numerous applications in, for example, growth accounting and productivity measurement and more generally when quality indices of labor input are required.

The results show that going from four levels of education to 29 levels and types of education would only entail a reduction in the average annual rate of growth of the index of labor services from 0.530 percentage points to 0.506 percentage points. Thus, the gain in precision gained from including more detail on education seems to be small.

The results found for Denmark in this paper are likely to essentially carry over to other developed countries, especially other continental European countries. This viewpoint is based on the observation that the wage structure across developed economies seems to be relatively similar, see Appendix A1.2. Thus, also elsewhere, we would expect that relatively little would be gained from including more educational detail than a few education levels.

APPENDIX

A1.1. Data

The data employed originate from the IDA database of Statistics Denmark and cover the period 1980–98. They encompass the entire Danish population cross-classified by a number of variables and give, for each group, *inter alia*, the number of people, the number of people with non-zero income information and the sum of incomes for people with non-zero incomes.

The data are cross-classified by the following variables.

- Employment status: nine groups, including self-employment, salaried employment, and pension.
- Job category: four employment statuses: employer, self-employed, co-working spouse, salaried employee.
- Sex.
- Age: eight groups.
- Industry: 126 industries, including 105 industries in the private sector and an auxiliary industry for non-employment. The 105 private sector industries include 55 industries within manufacturing.
- Education: 29 education types as combinations of level of education with type. By level we refer to the length of the education. The data distinguish five levels of education: short, medium and long further education, vocational education and a group of education levels with no formal skills. Each level of skilled education is further subdivided by type of education with eight groups: technical, natural sciences, social sciences, humanities, social and health, trade and secretarial, service, and other. A share of people born before 1921 are recorded with unknown education. After 1985 this group no longer appears in the data.

  Not all combinations of level and type exist and, for example, the

education levels with no formal skills are all classified as the education type "other."

- Working hours: nine groups based on pension payments and type of unemployment insurance.

This divides the population into a very large number of groups. The industry dimension has been omitted for the current analysis.

Employment has been converted to full-time equivalents based on the classification after working hours and adjusted to conform to the employment figures in the national accounts. No attempt has been made to convert to hours worked.

The income information relates to a whole year while employment and industry association relates to a certain date each year. Thus, a bias is introduced by some workers only working part of the year. If they happen to be unemployed on the date of measurement, their income is not attributed to an industry. Conversely, if they happen to be employed on the date of measurement, their income appears to be too low. In order to correct for this, the population is grouped by sex, age and education. For each group, the wage income of those employed on the date of measurement is adjusted to include the income located within the group of unemployed on the date of measurement. Thus, the wage income of those recorded as unemployed is distributed on industries proportionally to the wage income recorded within each industry. The correction is performed separately for each combination of sex, age and education group.

Generally, no wage income is recorded for employers, self-employed and coworking spouses. An income for these groups has been imputed from the wage income of salaried employees using a classification based on industry, sex, age, and education.

*A*1.2. *Education and labor market institutional setting in Denmark*

In Denmark, nine years of elementary school beginning at the age of 7 is compulsory with an optional 10th year. Unknown education and up to 10 years of elementary school is labelled as elementary schooling in our aggregate grouping of educations. After elementary school, pupils can go on to vocational training, in vocational schools or in an apprentice system. Alternatively, pupils can go on to various types of high school, typically for three years. High school is required to go on to further education, either at universities where typically a long further education is obtained after five years, or to professional schools for 3–4 years of medium further education. Medium and long further education is roughly equivalent to 4+ years of American college.

The Danish labor market is characterized by a high degree of unionization, in line with other Nordic countries. In Calmfors (1990), union models of wage formation have in fact been applied empirically with a reasonable degree of success to explain wage formation in the Nordic countries. The share of the Danish labor force and relative wages by four aggregate education groups for 1998 are presented in Table A1. The relative wages for the four broadly defined educational groups are presented in the table. It is seen that persons with further education have a wage income that exceeds that of persons with elementary school only by 41 percent on average. Furthermore, persons with high school and

TABLE A1

COMPOSITION OF LABOR FORCE AND RELATIVE WAGES, 1998

Aggregate Education Category	Share of Population	Relative Wage
Elementary	32.5	1.00
High school	6.2	1.18
Vocational	44.5	1.14
Further	16.8	1.41

Source: Statistics Denmark (2000) and own calculations.

vocational education earn wage income that exceeds that of the base group by about 18 percent.

An important issue in relation to the present study is whether the Danish wage structure over educational groups is more compressed than for other developed countries. This has been examined by, for example, Machin and Van Reenen (1998). In the comparison, skills are measured by grouping of labor after production and non-production workers. It is found that the Danish relative wages in manufacturing are similar to those in Japan, Sweden, the United Kingdom and the United States over the period 1973–89. Hence, the Danish wage structure does not seem to be more compressed than for other developed countries, if the wage structure in manufacturing indicates that of the total economy. For continental Europe, relative wages of unskilled labor have been more stable compared to the U.K. and the U.S. over the past two decades. For this reason, the main point established in this paper may hold also for other continental European countries.

REFERENCES

Calmfors, L. (ed.), *Wage Formation and Macroeconomic Policy in the Nordic Countries*, SNS and Oxford University Press, 1990.

Gu, W., F. C. Lee, and J. Tang, "Economic and productivity growth in Canadian industries," in D. W. Jorgenson and F. C. Lee (eds), *Industry-level Productivity and International Competitiveness between Canada and the United States*, Industry Canada Research Monograph, Ottawa, 77–120, 2001.

Ho, M. S. and D. W. Jorgenson, "The Quality of the U.S. Workforce," manuscript, Harvard University, 2001.

Jorgenson, D. W., "Information Technology and the U.S. Economy," *American Economic Review*, 91, 1–32, March 2001.

Jorgenson, D. W., F. M. Gollop, and B. M. Fraumeni, *Productivity and U.S. Economic Growth*, Harvard University Press, Boston, 1987.

Jorgenson, D. W. and K. J. Stiroh, "Raising the Speed Limit: U.S. Economic Growth in the Information Age," *Brookings Papers on Economic Activity*, 125–211, 2000.

Machin, S. and J. Van Reenen, "Technology and Changes in Skill Structure: Evidence from Seven OECD Countries," *Quarterly Journal of Economics*, 113(4), 1215–45, November 1998.

OECD, *Productivity Manual*, OECD, Paris, 2001

Oliner, S. D. and D. E. Sichel, "The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?," *Journal of Economic Perspectives*, 14, 3–22, Fall 2000.

Statistics Denmark, *IDA Database*, Statistics Denmark, Copenhagen, 2000.

Timmer, M., "Towards Productivity Comparisons Using the KLEMS Approach: An Overview of Sources and Methods, The KLEMS Project," Feasibility Studies, The Conference Board, 2000.