

## PART 2: EDUCATION SERVICES

1. The Eurostat handbook on prices and volume in national accounts has already provided a theoretical framework for non-market education services with stratification and quantity indicators, but little practical guidance for explicit quality adjustment, where several possibilities are reviewed without hierarchies and formulas. The lone differentiation of “products” seems to us insufficient in order to reflect the ever changing programs, pedagogies or quantities of inputs devoted to one pupil. The aim of this chapter 2 is therefore to develop the theoretical framework of the Eurostat handbook for education and to define subsequently formulas for explicit quality adjustment, by kind of education services, with some variants adapted to the various national information systems.

2. Consistently with chapter 1 of the present OECD handbook, in absence of relevant prices and “true” markets, we have especially explored outcome-based quality adjustments, which were already in the spirit of the Eurostat handbook (“*The quality of the output lies in its results, i.e. in the outcome.*”), with these refinements that direct outcomes have to be privileged on indirect outcomes, and that “the contribution of the producer to the change in the outcome” of the consumer is the straightforward measure of the quality adjustment. Such quality adjusted measures of volume seem to us superior to non quality adjusted price indices, so that contrarily to the Eurostat handbook we suggest to treat market activities with these methods initially developed for non-market activities, or to treat market and non-market activities together as a whole.

3. Section 1.7 has proved the similarity of the temporal and spatial notions of volume and price. Subsequently, we have tried to develop as much as possible a first large section common to these two issues, and we have proposed an application for the PPP exercise prior to precise recommendations for time series. Indeed, what is possible for international comparisons should be a basic starting point for any national time series, which can improve the international framework with their specific national organization and databases. Hence, this chapter progresses from generic to specific and adopts an unusual order for national accountants.

### 2.1. Terminology and concepts in education services

#### 2.1.1. What are education services?

##### 2.1.1.1. Dictionaries and ISCED-97 definitions

4. The Oxford Advanced Learner’s Dictionary (OALD) defines “education” as “*a process of teaching, training and learning, especially in schools or colleges, to improve knowledge and develop skills.*” The American Heritage dictionary provides two definitions: “*1. The act or process of educating or being educated. 2. The knowledge or skill obtained or developed by a learning process.*” These definitions outline the substance of education output, relying on a process (activity): teaching, training and learning, and contributing to an outcome (welfare status): knowledge and skills.

5. The UNESCO General Conference adopted a more precise definition of “education” at its 29<sup>th</sup> session in November 1997, consistent with its International Standard Classification of Education (ISCED-97): “*organised and sustained communication designed to bring about learning*”, where:

- “*Communication*” involves the transfer of information (messages, ideas, knowledge, strategies, etc.) between two or more persons;
- “*Organised*” communication is that which is planned in a pattern or sequence, with established aims or curricula. It should involve an educational agency that organises the learning situation and/or teachers who are employed (including unpaid volunteers) to consciously organise the communication;

- “*Sustained*” communication is that which has the elements of duration and continuity as part of the learning experience;
- “*Learning*” is taken as any change in behaviour, information, knowledge, understanding, attitudes, skills, or capabilities which can be retained and cannot be ascribed to physical growth or to the development of inherited behaviour patterns.

6. The contribution of this ISCED definition is that the “transfer of knowledge and skills” (“*communication*” = “transfer”... “*learning*” = “change in [...] knowledge and skills”) is “*organised and sustained*” by an economic unit which is not necessarily a regular school, with the help of curricula developed on several years. This leads to a classification of education levels defined in the same 29<sup>th</sup> session of the UNESCO General Conference and is also consistent with the scope of national accounts.

**Table 2.1.: International Standard Classification of Education – 1997, at one digit**

Level 0	Pre-primary education
Level 1	Primary education or first stage of basic education
Level 2	Lower secondary or second stage of basic education
Level 3	(Upper) secondary education
Level 4	Post-secondary non-tertiary education
Level 5	First stage of tertiary education
Level 6	Second stage of tertiary education

ISCED-97 has also defined 25 “fields of education” (see section 2.2.).

2.1.1.2. The content in activities and products for national accounts

7. The classification in the national accounts, which is one of the purposes of this handbook, is determined by the content of the ISIC<sup>1</sup> classification of activities and by the content of the CPC<sup>2</sup> products. The latter use the ISCED contents in their explanatory notes: As ISIC rev 4 is under construction (to be published in 2007) and CPC will follow, we can only provide the outlines of what will become the reference scope for education in national accounts and their relation with ISCED:

**Table 2.2.: Scope and classification of education according to upcoming ISIC and ISCED-97**

	ISIC rev 4 classes	ISCED-97 levels of education
Formal	8510 Pre-primary and primary education	Levels 0 and 1
	8521 General secondary education	Levels 2 and 3 oriented general
	8522 Technical and vocational secondary education	Levels 2 and 3 oriented vocational and technical
	8530 Higher education	Levels 4, 5 and 6
Informal	8541 Sports and recreation education	Not classified in ISCED-97 levels of education
	8542 Cultural education	
	8549 Other education n.e.c.	
Support	8550 Educational support activities	Not explicitly mentioned in ISCED-97 levels of education

<sup>1</sup> International Standard Industrial Classification of All Economic Activities, see chapter 1, section 1.1.1. and annex 2 for a detailed extract of ISIC rev 4 on education.

<sup>2</sup> Central Product Classification, see more information on <http://unstats.un.org/unsd/cr/family2.asp?Cl=16>

8. This tri-partition of education activities in (1) formal education, (2) informal education (3) support activities can be considered as the main structure of any comprehensive approach of education, and is more and more explicit with successive versions of international classifications. The formal education is the true focus of our handbook, consistent with the “education data” structured and collected internationally.

9. The informal education was usually not considered as education activity, but as sport or recreation activity. We will not provide an explicit guidance for the fields of this part (golf, bridge, music...), although they contribute also to GDP and are to be measured in volume and price, but the principles developed for formal education could also inspire them.

10. Educational support activities were even less considered as “education activities” in previous classifications. The true content of this ISIC class is still to be precised, we can assume it should content activities performed by educational units which are not “core educational activities” but which have no sense without the purpose of education and the context of the education units (otherwise they should be considered as secondary activities in accommodation, catering, transportation, inspection, etc.). The case of school bus transportation is the more typical for the distinction between activities / products on one side, “functions” on another side. In the classical view, its purpose / function is “education” in classifications like COFOG, COPNI or COICOP<sup>3</sup>, but its activity is “transportation”. Nevertheless, because the American school buses are very specific and paid by the schools, the North American Industry Classification System (NAICS) has recorded explicitly “School bus attendant services” in group 6117: “Educational Support Services”, inside the division 61: “Educational Services”. It seems to have inspired the new ISIC rev 4 class “8550: education support activities”. Because of their special nature, their only indirect impact on education outcome and their lack of comparability across nations, guidelines for calculation volume output of these activities are only slightly considered in this handbook.

**Table 2.3.: Functions of COFOG and COPNI in relation with ISCED-97**

	ISIC activities	COFOG	COPNI	ISCED-97
Formal education	Education: 8510	09.1: Pre-primary and primary education	04.1: Pre-primary and primary education	Levels 0 and 1
	Education: 8520	09.2: Secondary education	04.2: Secondary education	Levels 2 and 3
	Education: 8530	09.3: Post-secondary non-tertiary education	04.3: Post-secondary non-tertiary education	Level 4
09.4: Tertiary education		04.4: Tertiary education	Levels 5 and 6	
Informal	Education: 8540	09.5: Education not definable by level (of which adult education)	04.5: Education not definable by level (of which adult education)	Not classified by ISCED
Support	Education: 8550	09.6: Subsidiary services to education		Not classified by ISCED-97
	Research: 72	09.7: R&D Education	04.6: R&D Education	Not classified by ISCED-97
	Administration: 8412	09.8: Education n.e.c.	04.7: other educational services	Not classified by ISCED-97

<sup>3</sup> Classification of Individual Consumption According to Purpose, see chapter 1, annex 1 and more information on <http://unstats.un.org/unsd/cr/family2.asp?Cl=5>

11. The COFOG<sup>4</sup> concept of education is still broader than the scope of education activities and products in national accounts: research is included in COFOG/COPNI<sup>5</sup>, not in ISIC/CPC, and administration regulating the activities of education services is explicitly included in COFOG/COPNI but somewhere else in ISIC/CPC.

12. The only other noticeable difference between COFOG, COPNI on one hand, and ISIC rev 4 on the other hand, is the distinction of post-secondary non tertiary education within higher education. Previous version of ISIC was less consistent with COFOG.

13. This handbook will focus on the formal education, which are the levels of education explicitly defined in ISCED-97 classification, corresponding also to the ISIC activities 8510 to 8530, and to the first four items only of COFOG and COPNI “education” functions.

### **2.1.2. Inputs, processes, output and outcomes in education**

#### 2.1.2.1. An overview of inputs, processes, outputs and outcomes

14. The differentiation of inputs, activities/processes, output, direct and indirect outcomes is important in order to better delimitate what makes the substance of the output. As explained in Chapter 1, in absence of relevant prices and “true” market, and under the assumption of a one-dimensional utility, the quality component of the output can be defined as *the contribution of the producer of (education) services to the change in the (education) outcome of the customers* (here the pupils). It is consistent with the notion of “transfer of knowledge and skills towards the pupils” which has been mentioned in the dictionaries and the ISCED definitions.

15. The previous methods used in national accounts were relying on deflated inputs: deflated wages of teachers plus deflated intermediate consumption plus deflated fixed capital consumption (depreciation of buildings for schools and universities). Because of the importance of wages in this sum, this could be dubbed as a method whereby the volume output was correlated to the number of teachers (deflated wages are correlated to number of teachers). The method proposed in the present handbook voluntarily moves away from this approach with the objective of measuring the impact of education services as the amount of skills and knowledge transferred to the customers (the pupils).

16. The pupil-hours, 1<sup>st</sup> best quantity index of the output according to the Eurostat handbook on volume and price, are properly speaking a process indicator.

17. The human capital theory relies on an indirect outcome (future real earnings). The literature on the human capital theory is not always quite clear with the impact of environmental factors, to be neutralized or not in a strict notion of “output”.

18. The incremental contribution of “knowledge” to the growth of GDP (net of environmental factors) still relies on an indirect outcome and is not consistent with a notion of “individual services”. However, this idea of an increasing content of education should be expressed by some way in the quality indicator, which lone scores will usually miss (because usually based on the past content only).

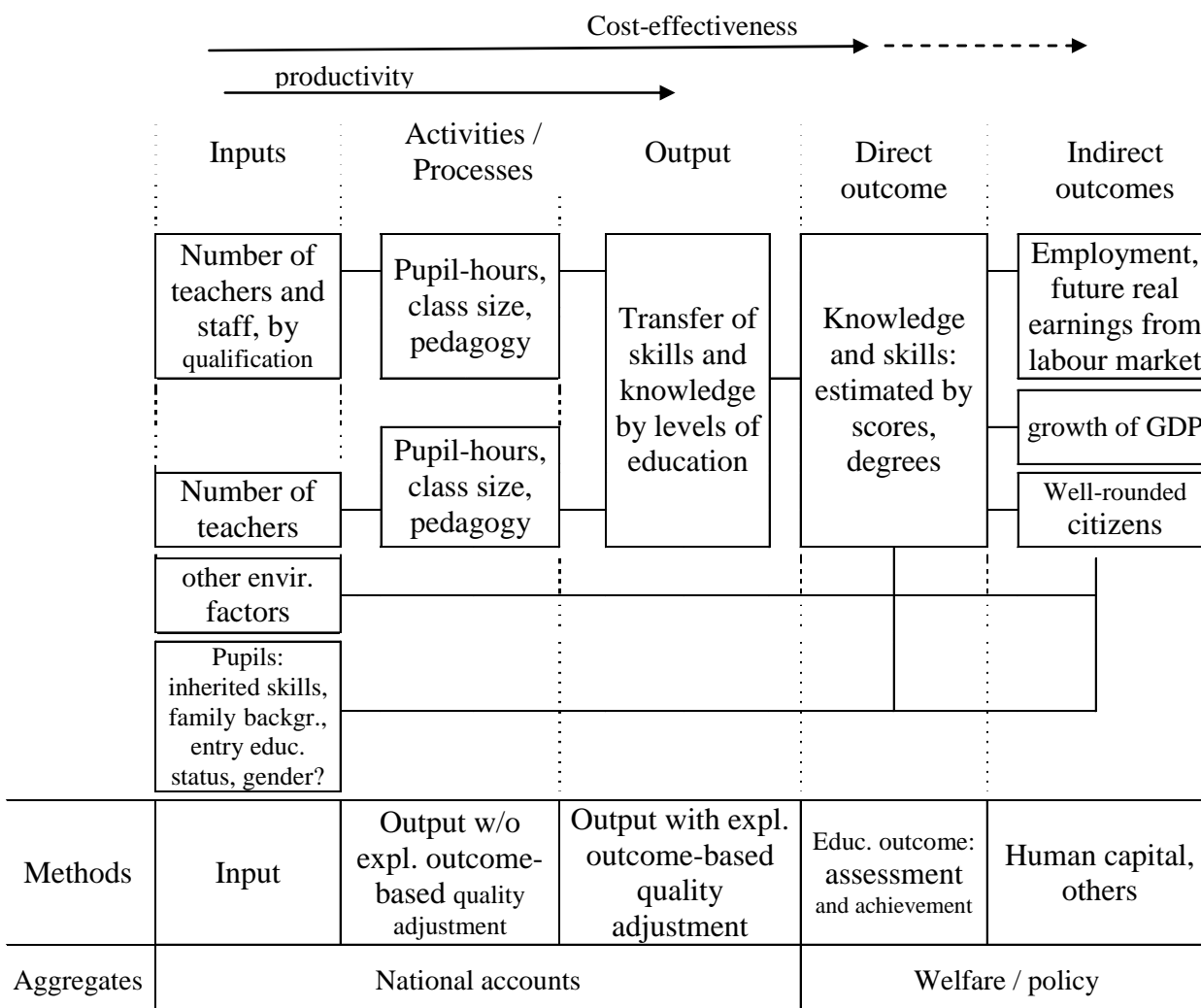
---

<sup>4</sup> Classification of the Functions of Government, see chapter 1, annex 1 and more information on <http://unstats.un.org/unsd/cr/family2.asp?Cl=4>

<sup>5</sup> Classification of the Purposes of Non-Profit Institutions Serving Households, see more information on <http://unstats.un.org/unsd/cr/family2.asp?Cl=6>

19. Our approach implies to define first what the direct outcome indicators of education are, and, second, to model what the measures of the output can be inside these outcome indicators, because we want a quality adjustment strictly limited to the change in the outcome that is attributable to the producer.

**Scheme 2.1.: Articulation between inputs, outputs and outcomes in education**



2.1.2.2. The education outcomes

20. The direct educational outcome for individuals is their education status, i.e. their level of knowledge and skills.

We propose to use academic scores as the usual measurement of the level of knowledge and skills.

Part of knowledge and skills owned by individuals has been inherited (genetic) or acquired by the pupils and students themselves, part has been assimilated with the help of their family background, and at last part has been transferred by the education economic units.

21. For higher education, the direct outcome may be summarised by the delivery of a certain degree or diploma, but this latter provides generally from (sometimes implicit) academic scores above a given threshold. The attainment of a specific employment in line with a vocational curriculum, or the labour income connected to a certain amount of knowledge and skills can also be seen as concrete results, but they

are only indirect outcomes of the education process. The most direct outcome of education is better represented by scores on academic scales.

22. The main direct educational outcome for a community, its overall educational status, is the sum of all individual levels of knowledge and skills.

Specific aspects of the distribution of individual levels of knowledge and skills can define complementary information for the education status of a community and other education direct outcomes for the community, important for the political goals assigned to the education system (inequality measured by Standard Deviation of scores, excellence measured by part of pupils performing more than certain thresholds of scores, illiteracy or school failure measured by part of pupils performing less than certain thresholds of scores, etc.). These complementary characteristics of the direct outcome for the community can be expressed in the same measurement unit as the main direct outcome, although they are generally expressed in terms of percentage of this latter. These important secondary dimensions of the direct outcome for the political goals and the assessment of public services efficiency do not match the classical philosophy of national accounts for an individual service, which sees the community as a simple addition of individuals without any consideration of “inequality”.

23. In formal education, academic scores can usually be defined and harmonized at a fine level of education fields and grades. Common conventions can be used to weight these elementary scores to obtain a multidisciplinary score for a grade. A same education field can also be measured and compared over different grades. But in practice, a common academic score rarely exists for all education fields and grades.

24. However, all academic scores are subjective. Each teacher has a different subjective method of scoring. The rankings of pupils and the respect of some minimal thresholds can be reputed “objective”, but how can we assess that such a pupil has knowledge and skills for twice much than another one? The first elementary step usually consists in normalising the Standard Deviation of scores. It seems that many national and international assessments are standardized with two alternative conventions: either 1 Standard Deviation = 10% of the average (American NAEP mathematic test), either 1 SD = 20% of the average (PISA, PIRLS and TIMSS, all international tests). If we suppose that the distribution of scores respects a “normal” frame, the 16<sup>th</sup> centile of pupils performs the average – 1 SD, while the 84<sup>th</sup> centile of pupils scores the average + 1 SD. It means, in case of 1 SD = 20%, that the second category is reputed knowing and skilled  $(1 + 0.2) / (1 - 0.2) = 1.5$  times much than the first category, or in case of 1 SD = 10%, that the second is worth  $(1 + 0.1) / (1 - 0.1) = 1.22$  more than the first one on the academic scale. International or temporal comparisons will therefore command using the same norm. Scores realised under a certain SD can be approximated in another SD by simple mathematical (affine) transformation.

25. However, even using a normalised SD, the question remains whether this normalised scale can be corroborated with other measures. In context of economic indicators such as national accounts, it seems natural to correlate the scale for knowledge and skills and future real earnings. Alan KRUEGER, in “Economic considerations and class size” (*Economic Journal*, 2003), has reviewed three studies:

- “MURNANE, WILET and LEVY (1995) estimate that male high school seniors who scored one standard deviation (SD) higher on the basic math achievement test in 1980 earned 7,7 percent higher earnings six years later, based on data from the High School and Beyond survey. The comparable figures for females was 10,9 percent.” ;

- “CURRIE and THOMAS (1999) use the British National Child Development Study to examine the relationship between math and reading test scores at age 7 and earnings at age 33. They find that students who score in the upper quartile of the reading exam earn 20 percent more than students who score in the lower quartile of the exam, while students in the top quartile of the math exam earn another 19 percent one. Assuming normality, the average student in the top quartile scores about 2,5 standard deviations

higher than the average student in the bottom quartile, so their results imply that one SD increase in reading test performance is associated with 8,0 percent higher earnings, while a one standard deviation increase in the math test is associated with 7,6 percent higher earnings” ;

- “NEAL and JOHNSON (1996) use the National Longitudinal Survey of Youth to estimate the effect of students’ scores on the Armed Forces Qualification Test (AFQT) taken at ages 15-18 (adjusted for age when the test was taken) on their earnings at age 26-29. They find that a one SD increase in scores is associated with about 20 percent higher earnings for both men and women”.

If we assume the hypothesis of proportionality between quantity of knowledge and skills and future real earnings, these studies suggest either a relative standard deviation for education outcomes about 8% (first two studies), either about 20% (last study). This is quite a difference, but it is the same as in academic scores. Hence, we cannot conclude if 20% is more “objective” or “relevant” than 10%, but these two options are the most common ones for both direct and indirect outcomes. We conclude provisionally that pure academic scales for direct outcomes are necessarily subjective, but they can be conventionally standardized with either a relative standard deviation = 10%, either a relative standard deviation = 20%.

### 2.1.2.3. The education output

26. The main quality component of the output of an economic unit for a given pupil is the expected/average transfer of knowledge and skills delivered by this education unit towards this pupil, during a certain period of time, by level of education and curriculum.

The main quality component of the output of an education system is the sum of the expected transfers of knowledge and skills towards all its pupils, i.e. the total transfer of knowledge and skills it performs.

These are the proper translations for education services of the more general definition in non-market services that the quality of the output lies in the expected *contribution* of the producer to the change in the outcome of the consumer, by kind of services.

27. In case the outcome can be assessed by a score, the main quality component of the output is homogeneous to a change in the score. The basic definition of the output of school consumed by a pupil is proportional to the contribution to the change in the score obtained by the pupils of the same class between an examination conducted at the beginning of the school year and the same examination conducted at the end of the school year, everything else being equal (without contribution of environmental factors).

28. *Contribution* of the education unit mean that only the part that is strictly attributable to the education unit in the change in the education outcome of this pupil should be included. Other contributions should be excluded, such as:

- pupil contribution itself (inherited skills or special efforts, motivation);
- family contribution;
- another education unit contribution – extra school remedial courses, private tutoring if we focus on a unit and not on the whole education system..

These other contributions to the pupils’ level of knowledge and skills explain mainly why different pupils obtain different outcomes by receiving the same teaching (hence, the same output).

29. *Expected*, which can be more intuitively replaced by *average*, means consistently from the point of view of one consumer, that the output and the actual final consumption are not determined by each individual pupil’s characteristics and ability to derive a utility from the service provided, but they are influenced by the general distribution of these characteristics (basically the average), which provide the

market equilibrium with the user value side of the price. This is in line with the recommendations of the SNA 1993, not to confuse the notion of volume with the actual benefit or utility to be individually derived from its consumption. This makes also fit the definition of this handbook “*expected/average transfer of knowledge and skills delivered by an education unit or by the education system towards its pupils*” with the definition of the European handbook: “*In the area of education, the output can be defined as the quantity of teaching (that is, the transfer of knowledge, successfully or not) for a pupil, whereas the related outcomes are the skill and knowledge that a pupil achieves.*” The transfer can be unsuccessful for an individual pupil, the output is reputed identical for all pupils of a same class receiving the same teaching, therefore the measurement of the output has to be made “in expectation” relatively to one pupil, we can assimilate to “in average” in case of large numbers. For a whole class or a whole nation, the total expected transfers equal the total actual transfers.

30. If we combine a quantity component that is the number of enrolled pupils and the main quality component of the education output we have defined above, we can consider the following equalities, by level of education and curriculum:

$$\begin{aligned}
 \text{Education\_output} &= N (\text{number of pupils}) * (\text{average\_change\_in\_score\_due\_to\_school}) \\
 &= N (\text{number of pupils}) * (\text{average\_transfer\_of\_knowledge\_due\_to\_school}) \\
 &= \sum_{i=1}^N (\text{average\_transfer\_of\_knowledge}) = \sum_{i=1}^N (\text{expected\_transfer\_of\_knowledge})
 \end{aligned}$$

31. This set of equalities makes sense with a number of pupils, but would not with pupil-hours. Indeed, it would lead to a double accounting. Imagine a pupil with an attendance of 90% instead of 100%. As he has received 10% less teaching, it is probable that he will be punished by obtaining around 90% of the change in score relatively to his “normal” schoolmates during this year. If we cumulate the quantity effect (pupil hours) and the quality indicator (change in scores) as in the above formula, we will obtain 81% instead of 100% vis-à-vis his assiduous schoolmates, but it is incorrect, for the negative factors have been counted twice, and the transfer of knowledge can be assumed 90%.

32. Our formula is fully consistent with the European handbook:

- if no indicator of change in scores is available, and no model is built for this part, this formula provides basically the number of pupils, which is the second best quantity indicator recommended by the European handbook;
- if no indicator of change in scores is available, but the process of transfer of knowledge is reputed continuous and proportional to pupil-hours, we can take into account this proportionality in the quality indicator and in this case the formula provides the number of pupil-hours, which is the first best quantity indicator recommended by the European handbook;
- If there are data on scores, the quality adjustment provided by the change in score in a “value added” view (end of the year/curriculum – beginning of the year/curriculum) corresponds to the recommendations of the Eurostat handbook: “*The quality of the output lies in its results, i.e. in the outcome.*” “*For examples of outcome-based indicators it may be appropriate to consider whether examination data (perhaps with a "value added" approach, where pupil abilities are measured before and after a period of schooling) [...] can be used to adjust for changes in quality.*”

#### *Change in scores versus absolute levels*

33. As explained in our basic definition of output, we want the quality adjustment factor to intervene as proportional to the *change* in academic scores between the beginning and the end of the school year, everything else being equal. Unfortunately, there are, in practice, almost no academic scoring systems that

give such results, as exams are not comparable between different ages. There exist however some cases, like the American NAEP math test, quoted by Michael CHRISTIAN in its contribution to the London workshop of October 2006, or the UK “4 key stages”. These examinations take place at different ages and seem to be comparable between ages, thus allow deriving a change in scores, however over a number of years. On the NAEP, M. Christian says:

*“Analysis of National Assessment of Educational Progress (NAEP) math test scores suggests that a standard deviation of test scores is the equivalent of 3.3 years of schooling (the standard deviation of math NAEP scores for 17-year-olds is about 31 points; this was approximated by observing the percentile distribution of scores in 1996 and assuming a normal distribution. The average math NAEP score improved from 231 at age 9 to 307 at age 17. Dividing the difference between these two scores by 8 yields an annual NAEP gain of 9.5 points, which is about 1/3.3 the cross-sectional standard deviation of 31).”*

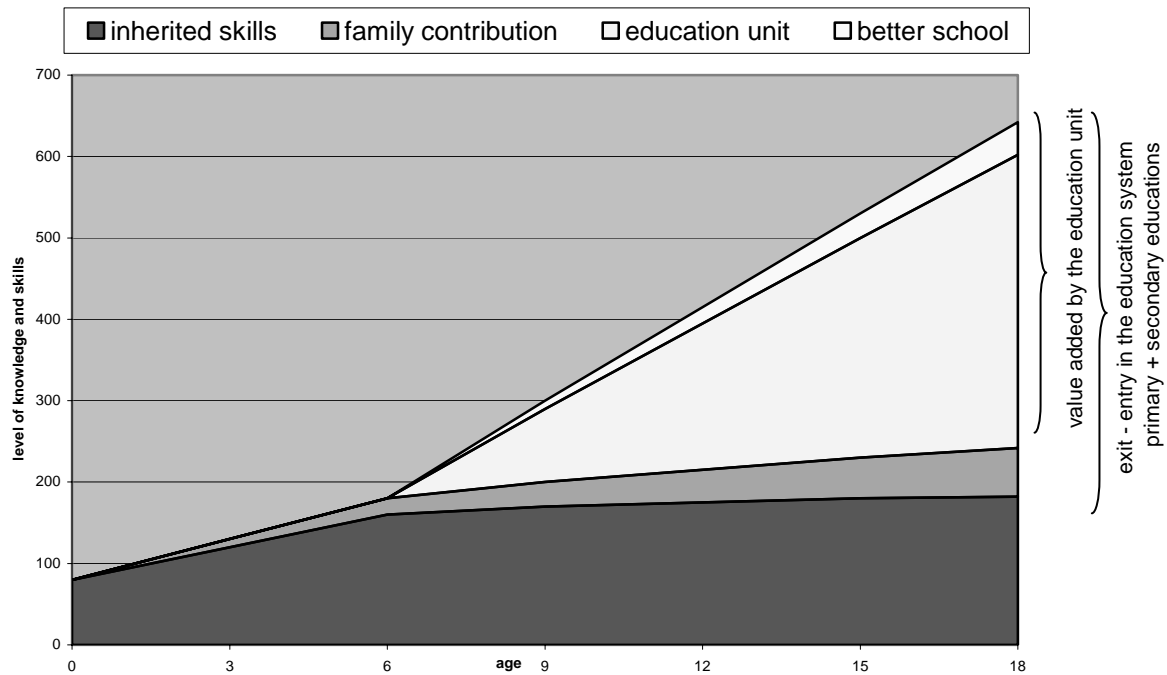
34. Thus a relation can be made between scores at different ages, however over several years. The importance of measuring changes in scores rather than absolute level of scores can be illustrated by the following example, inspired by the NAEP test. If we suppose two schools of 100 pupils per grade between ages 9 and 17, one getting average scores (from 231 at 9 – the entry – to 307 at 17 – the end) and the other one getting the average plus a Standard Deviation in the final outcome with the same characteristics in the entry (from 231 at 9 to 338 at 17), equally affected by the inherited skills and the family background, the respective transfers of knowledge and skills and contributions to the change of the outcome for each pupil are valued 76 and 107 for 8 years, thus an impact of +41% (107/76) attributable to the fact that the second school is better. Thus the final outcome differs by 10% (31 on 307), but the quality adjustment provides a difference of 41% for the output.

#### *Exclusion of family and individual background*

35. We may consider that the education outcome is made from 3 contributions: (1) inherited skills, (2) knowledge and skills attributable to the family background, and (3) knowledge and skills transferred by the education units. As our objective is to measure the output of the education services, we have to measure the third contribution, separately from the two first. The graph below shows a typical situation whereby the level of knowledge and skills, measured in terms of academic scores appears on the vertical axis. As can be seen from the graph, the increase of the academic score for an individual can originate from inherited skills, family contribution, and the education unit. We would like to measure only the latter, which is the light colored part in the graph.

We will discuss in section 2.2 and 2.3 on how to neutralize these non wanted effects (inherited skills, family contribution, or a proxy which is the education status at the entry into the education system) in the practical context of cross-country comparisons and time series comparisons.

Chart 2.1 Education status with age



*Management of the time-lag between outcomes and outputs*

36. A major problem is the fact that national or international examinations occur at some points in time, for one or few grades, on samples rather than, as needed, every year for all pupils. A model may be necessary to link the global output of one year to the outcomes measured the further years. This will be discussed also in sections 2.2 and 2.3.

The European handbook was aware of this drawback of a long time-lag and multiple possible revisions: *“The quality of the output lies in its results, i.e. in the outcome.”* *“For examples of outcome-based indicators it may be appropriate to consider whether examination data (perhaps with a “value added” approach, where pupil abilities are measured before and after a period of schooling) or “moving up” data (that is the proportion of pupils who pass through to the higher year) can be used to adjust for changes in quality.”* *“A problem with this approach is the possible time lag between a change in the quality of the output and a change in a particular outcome indicator.”*

Anyway, the management of revisions of national accounts series with continuous new informations relevant for past years is a classical paradigm. It can be admitted that current revisions occur during three years, and then a new benchmark / revision of long series has to be waited for, in order to introduce all the new information accumulated after the delay of three years.

*Measuring at given years or at given grades?*

37. The remark of the Eurostat handbook on the number of pupils moving up deserves a comment. If scores “at the entry” and “at the exit” of the education system are measured at given ages, this qualitative indicator is consistent with a quantitative indicator relying on the number of enrolled pupils. If these scores are measured at given grades, this qualitative indicator is consistent with a quantitative indicator relying on the number of pupils moving up:

$$\begin{aligned} \text{transfer of knowledge} &= \text{number\_pupils\_getting\_1\_year\_older} * \text{average\_transfer\_1\_year} \\ &= \text{number\_pupils\_moving\_up\_1\_grade} * \text{average\_transfer\_1\_grade} \end{aligned}$$

*Non-educational goods and services included in “education” in national accounts*

38. All the reasoning on the education output and the formulas which could summarize it were up to now based on the “core educational services”, i.e. on the transfer of skills and knowledge. Non-educational goods and services linked to education (catering, boarding services, transport services) should be deflated or directly estimated in volume apart, consistently with each kind of product provided, as it is recommended in the Eurostat handbook (“*For the purposes of comparability, it is important that any non-education services provided alongside educational services (for example school meals, transport and research in universities) are separately identified and deflated using appropriate price indices where these are available [...]*”):

$$\begin{aligned} \text{Education output} &= N(\text{number\_of\_pupils}) * (\text{average\_transfer\_of\_knowledge}) \\ &+ \text{non-educational linked services.} \end{aligned}$$

But it can happen that this amount is only estimated as a whole, without clear distinction by products and knowledge of relevant prices. In this case, the right complete formula could be, as a second best:

$$\begin{aligned} \text{Education output} &= N(\text{number\_of\_pupils}) * (\text{average\_transfer\_of\_knowledge}) \\ &* \left( \frac{\text{total\_xpenditures}}{\text{core\_educational}} \right). \end{aligned}$$

#### 2.1.2.4. The education inputs

39. The labour input of teachers and auxiliary staff is the main component of the education services (often circa 85% of the total costs). In order to calculate an accurate productivity ratio, we would suggest dividing basically the labour input between “educational staff” and “auxiliary or non-educational staff”. The educational staff can be split between kindergarten teachers, primary teachers, secondary teachers and university lecturers, which only reflect the grades, but also between holders of academic titles and others (by “skill level”). The non-educational staff should also be divided by skill level. This is the theoretical input approach of education in the PPP.

40. In the intermediate consumption, a particular point of interest is to check if accommodation, catering, books or other pedagogical materials are described as being provided apart to the pupils (in final consumption of other products, which is the best practice of the SNA) or as being intermediately consumed and provided to the pupils in a package “education services” (which is often the view of the economic unit). If such a status is modified between two periods or two national systems, output needs a quality adjustment consistent with the perimeter of inputs.

41. The environmental factors of the consumers, also called non-discretionary inputs, should be neutralized in the estimate of the output from a measure of the outcome, as an “environmental bias”. In education, it means that the family background and even the inherited skills have to be subtracted from the education outcome to provide an estimate of the output. Similar ideas are often practiced, either in international assessments like PISA and its “Economic, Social, and Cultural Status (ESCS) correction” (all countries are extrapolated as if all had the same common ESCS structure as the average OECD, see more in section 2.2), either in national comparisons of education units (comparisons are made with “expected outcomes” counting in the economic and social characteristics of the pupils), but it is often forgotten in national time-series! The economic, social and cultural structure moves slowly, too much to be observed

with reliability each year, but this change has a strong influence over 10 or 20 years, and national accounts have also a purpose to provide consistent long series.

In theory, there are plenty of environmental factors such as epidemics, climatic calamities, transportation strikes... which can influence badly the outcomes, without that the producer efficiency is questioned, but the producer efficiency is another matter than the producer output. The problematic in national accounts and in this handbook is basically not to count for the output what is due to the consumer himself.

42. In case of extra-school remedial courses, which can be considered as a non discretionary input for the regular school system output, it could be useful to determine what is to be attributed to whom, but it is perhaps not necessary for the national aggregates in NA, which will assess first of all the volume of the whole industry (regular school system and private tutoring together). It is likely that remedial courses are more expensive (reported to gain in scores) than main education units, because of the general theory of decreasing output, and because of the higher ratio teachers / pupils they rely on (often 1 teacher for 1 pupil alone). Regression techniques could be used to isolate the impact of remedial courses / private tutoring.

### **2.1.3. Stratification, weights and formulas**

#### 2.1.3.1. Stratification

43. In section 2.1, we have seen that the different kinds of formal education services, according to all international classifications, were basically the ISCED-97 levels of education, and secondarily their orientation, or the fields of education. This justifies a consistent stratification for the calculation of the volume of output, according to the rules of national accounts and to the recommendations of the Eurostat handbook on price and volume in national accounts: *“the quantity of teaching received by the students, adjusted to allow for the qualities of the services provided, for each type of education“*, which is explicitly developed later in: *“stratification at least to the following categories - pre-school, primary, lower secondary, upper secondary (general/vocational), higher education (university/other), and other education. In the case of higher education courses, there should be stratification by subject (for example differentiating science and medical courses from arts courses).“*

44. A detailed stratification of the various education output is essential for several reasons. First, users do not want to have only the aggregate result for total education, but would like to be able to understand the contribution of the various levels, or even units. Second, as explained in sections 1.4 and 1.5, different purposes (thus different outcomes) lead to different products, and it is not appropriate to add units of each product without taking into account the relative “prices” of these units.

45. In the context of education it means that one cannot add a pupil of ISCED 1 to a student of ISCED 6. The “price” of these two “products” of different purposes is different, and should be used to weight the quantities. However, “price” is associated to market activities, which are marginal or at least represent a minority of education services. There is no “price” for pupils or students for schools or universities that are financed by government or non profit institutions and made free or quasi-free to students, but there are costs. As discussed in section 1.5, the choice of this handbook is to rely on costs to weight the various “products” (strata / groups of substitutable items) for non market services. In this context, the stratification used should be tailored to take as much into account the different products and associate to each of them its unit costs. These unit costs will be used as weights when aggregating number of pupils or students. By definition, this means that a student of ISCED 6 will weight more than a pupil of ISCED 1, as the former costs more than the latter. This is appropriate in the context of an economic approach of education.

46. The stratification can also require different kinds of quality adjustments. The basic index developed in section 2.1.2.3., which we have relied on academic scores, can have to be adapted or even replaced for some kinds of education services, levels of education or whatever.

47. The below table provides a minimal stratification, in the spirit of the now universal tri-partition of education services and in line with the ISCED-97 levels of education, but also a more suitable one, justified by an heterogeneity of purposes, hence of outcomes, which requires also a more detailed database.

**Table 2.6.: Minimal and more suitable stratifications  
(for both temporal and spatial purposes)**

Formal education	Level 0	Pre-primary education	All classes
	Level 1	Primary education or first stage of basic education	Normal classes or pupils
			Special classes or handicapped pupils
	Level 2	Lower secondary or second stage of basic education	Normal classes or pupils
			Special classes or handicapped pupils
	Level 3	Upper secondary education	General + pre-vocational
			Vocational
	Level 4	Post-secondary non-tertiary education	General / vocational if available
	Level 5B	More practical and occupationally specific programmes tertiary education	All classes or by professional purpose
	Level 5A + 6	More theoretically based programmes tertiary education	By fields of education and/or prestige of education unit, or by equivalences of degrees...
	Adult and other informal education	Adult and other education, anticipating extension of "education" content in ISIC rev 4, class 8540.	Adult general education
			Adult vocational education
			Computer training
			Driving lessons
Music lessons			
Other cultural and artistic lessons			
Sport lessons			
Recreational lessons			
Education support activities	According to what will be retained in class 8550 of ISIC rev.4	Other education activities	

This handbook focuses on formal education thus the part on adult and other informal education and education support activities appears only for information.

48. The additional sub-stratification between normal and “special” classes “for physically or mentally handicapped pupils”, as it is suggested by ISIC rev 4<sup>6</sup>, in primary and lower secondary educations is justified by the importance of a compulsory ancillary activity in the second case (social services), involving more costs (non educational staff, special equipments). If handicapped pupils are taught in normal classes (it is considered better for them), the same idea of extra social services provided to them will be simply reflected by the number of handicapped pupils weighted by an estimate of their relative unit costs.

49. The additional sub-stratification between general and pre-vocational or pre-technical on one hand, vocational or technical programmes on the other hand, within the upper secondary education and the post-secondary non tertiary level of education (if available, this last stratum is often small), is justified by their different purposes: general knowledge and skills in the first case, specific knowledge and skills oriented towards a specific employment in the third case.

Note that the use of cost weights can lead to a paradox at first sight for education experts, if vocational curricula are more expensive than general ones: the substitution of vocational curricula to general ones makes a positive contribution to GDP, although it would probably not reflect a higher transfer of knowledge and skills...

50. The sub-stratification of tertiary education is less obvious. It is desirable because of the strong heterogeneity of the institutions and the curricula, of unequal “prestige”. It is a common practice in international organizations to group “first stage: more theoretically based programmes (5A)” and “second stage: advanced research qualification (6)” on one hand, “first stage: more practical and occupationally specific programmes (5B)” on the other hand, in the same spirit as general and vocational programmes in the secondary education. Inside the first category, the Bologna process (1999) has defined for European countries an equivalence of degrees which provides a view of first cycle (ending by a bachelor’s degree) / second cycle (ending by a master’s degree or a doctorate) which is slightly different from the first stage / second stage distinction in ISCED-97, and which suggests to evaluate simultaneously these two levels of education around the three equivalences of degrees. The European handbook recommends a stratification relying on the distinction university / others and/or on the fields of education (note that ISCED-97 has also defined 25 fields of education, of which: “14-education”, “21-arts”, “22-humanities”... “86-security services”). In our mind, it is particularly relevant if this sub-stratification allows a consistent ranking of prestige and costs or if an obvious “user value” allows to assimilate one kind of study as the equivalent of the double of another one for instance.

*A finer stratification relying on processes is not always suitable*

51. While more detailed stratifications may be useful, it is not however always recommended to use any detailed stratification (see also box 1 in section 1.4: “Why narrow specifications of products may not always be sufficient to capture quality change”). In particular, a stratification which would cross cut the ISCED classification with characteristics such as pedagogical methods and/or class sizes could lead to the opposite result than the one expected. Such a “process-based” classification would assume that relative costs of processes are always consistent with relative “user values” for consumers, while a correct “output method” consists in measuring the latter directly. Consequently, this “process-based stratification” could lead to an index of volume equal to the old input based index, thus leading to zero productivity, by construction.

---

<sup>6</sup> This explanatory note of the draft ISIC rev 4 division for “education” provided to “handicapped pupils” is more restrictive than the OECD Handbook for Internationally Comparative Education Statistics, which gives a description of “special education” (p.30) : “*education provision for mentally, physically, or emotionally or socially disadvantaged.*”

52. To give a concrete example, let us imagine stratification by class size: each class size defines a different product. Let us suppose that costs are limited to the wage of teachers, the wage being unique for all teachers. The elementary index of quantity (*Iquant*) in this simplified “process-based stratification” is the change of the global number of pupils (*Inum\_pupils*) attending a class of size *i*, and the index of volume (*Ivol*) is the weighted average of these indices of quantity, the weights being the global costs for all the classes of the same size *i*. Under our simplifying assumption of wage uniqueness, we can assimilate these weights to the total number of teachers.

$$Ivol = \frac{\sum_{i=1}^N \text{cost}(i)(n-1) * Iquant(i)}{\sum_{i=1}^N \text{cost}(i)(n-1)}$$

$$\Rightarrow Ivol = \frac{\sum_{i=1}^N \text{num\_teachers}(i)(n-1) * Inum\_pupils(i)}{\sum_{i=1}^N \text{num\_teachers}(i)(n-1)}$$

As in each stratum, defined by class size, the number of teachers and the number of pupils are proportional:  $\text{num\_teachers}(i) * i = \text{num\_pupils}(i) \Rightarrow I\text{num\_teachers}(i) = I\text{num\_pupils}(i)$

$$\Rightarrow Ivol = \frac{\sum_{i=1}^N \text{num\_teachers}(i)(n)}{\text{num\_teachers}(n-1)}$$

$$\Rightarrow Ivol = \text{num\_teachers}(n) / \text{num\_teachers}(n-1) = I \text{ num\_teachers}$$

Thus the volume index obtained by this “process-based stratification” is the change in the number of teachers, i.e. the same volume index as in the traditional input method.

#### 2.1.3.2. Alternative quantity and quality indicators

53. The basic formula, for the strict education part of the education output, is:

$$\text{Education output} = N (\text{number\_of\_pupils}) * (\text{average\_transfer\_of\_knowledge})$$

Our basic assumption is that the transfer of knowledge can be measured by a change in scores attributable to the producer. However, scores at exams are not always the only and most relevant outcome-based indicators.

54. In pre-primary schools, scores at exams are almost inexistent, as any evaluation of the education status at the end of the curriculum. Pupil-hours would then be a more relevant quantity indicators as they reflect better the service provided to the parents of the pupils (as kindergartens) and also the intensity of the education provided to the pupils (as pre-primary education).

55. In vocational or technical upper secondary education, in post-secondary non-tertiary education, or in vocational tertiary education, the professional insertion of pupils in an employment corresponding to their qualification could be considered as an adequate measure of outcome, although indirect. If no indicator on scores on a vocational specific scale is available, the quality indicator could be the proportion of pupils finding a job in the year following the end of the curriculum, in the corresponding qualification of employment.

#### *The case of tertiary education*

56. In tertiary education, the achievement of a degree is the more common outcome indicator, which relies on the respect of a threshold on an academic scale, but the individual scores are rarely recorded. The European handbook admits the use of this alternative quantity indicator in tertiary education, or as a basis

for quality adjustment: “numbers of students as a quantity indicator [...] is unrelated to the performance of the schools. An indicator such as the number of graduates is – to a certain extent – influenced by the quality of the teaching, and would therefore provide an indicator for quality-adjusted output.” We can note some problems, such as the estimates of the relative values of the various degrees, the double diplomas awarding a unique curriculum (especially with international curricula), and the fact that the contribution of the university is not separated from the entry education status.

57. In Europe as it has been said above (see section 2.1.3.1), the Bologna process has harmonized since 1999 the values of the various degrees in more theoretically based programmes of tertiary education (ISCED-97 levels 5A + 6) by classifying most of them in bachelor’s degrees (average 3 years of studies<sup>7</sup>), master’s degrees<sup>8</sup> (average extra 2 years of studies within second cycle but first stage) and doctorates (average 3 years of studies within second cycle second stage, ISCED-97 level 6). A combination of the number of degrees delivered, by kind of degrees (including number of students quitting tertiary education without degree), and by education field, could then be the right outcome-based quantity indicator. The degrees could be weighted by their relative “user values” within the product “tertiary education”. The typical numbers of years of studies could provide a basic “resource cost” view but estimates of average returns on future real earnings would be a better candidate.

58. In case of successive degrees, as it is the spirit of the Bologna process, the volume of tertiary education would become:

$I_{vol} = 1 * I_{quant} \text{ (number of students quitting w/o degree)}$ $+ b * I_{quant} \text{ (number of bachelor's degrees delivered in n)}$ $+ m * I_{quant} \text{ (number of master's degrees delivered in n)}$ $+ d * I_{quant} \text{ (number of doctorate's degrees delivered in n)}^9$
---

As the “values” of degrees are relative, “1” can be the value of the transfer of knowledge and skills provided to students quitting tertiary education without degrees (some students leave empty-handed after 1 year, or after 2 years, or after some months, in average little more than 1 year, and tests prove they have been transferred some knowledge and skills, recognized on the labour market). If b, m and d were taken equal respectively to 3, 2 and 3, the relative typical durations of studies, the volume of tertiary education would be very close to the simple number of enrolled students for each year<sup>10</sup>, easier to obtain but perhaps not exempt of double accounting (if a double degree awards a unique curriculum, do we have to record it as a double output?).

<sup>7</sup> 4 years in the United States.

<sup>8</sup> The UOE data collection records already the number of first degrees in ISCED 5A (usually bachelors) and second degrees (usually masters).

<sup>9</sup> This formula could be lagged such as in box 1 of section 2.2.

<sup>10</sup> An American experimental calculation has found a slight difference of 0.2% between a volume index relying on enrolled students (1.2% between 1980 and 2001) and a counting of degrees weighted by the number of years typically required to completion (1.4%). The difference comes from the varying rate of failure in the final exams (Michael CHRISTIAN, London workshop, October 2006).

59. If b, m and d were taken from the economic return on the labour market (contribution to future real earnings), some US experimental calculations<sup>11</sup> suggest that the value of one higher education year is decreasing: “Among white men, Jaeger and Page found that the total return to four years of undergraduate college is 17.8 percent and the additional return to two or more years of graduate school is 4.6 percent. This could be interpreted as a  $17.8 \div 4 = 4.45$  percent return to a year of undergraduate schooling and a  $4.6 \div 2 = 2.3$  percent return to a year of graduate schooling. Additionally, Jaeger and Page found that white men with occupational associate's degrees earned 0.7 percent less than those with some college but no degree, while those with academic associate's degrees earned 10.8 percent more and those with bachelor's degrees earned 16.2 percent more. If about half of associate's degrees are occupational and half are academic, this implies an average return to associate's degrees of  $(-0.7 + 10.8) \div 2 = 5.05$  percent and a return to a bachelor's degree of 16.2 percent. Finally, Jaeger and Page find returns of 5.0 percent to master's degrees, 28.6 percent to first-professional degrees, and 6.7 percent to doctoral degrees.” It would provide relative values of 4, 1.2 ( $5 \cdot 4 / 16.2$ ) and 1.7 ( $6.7 \cdot 4 / 16.2$ ) for b, m and d, instead of 4, 2 and 3 (standard durations in the US). The value of one higher education year via the contribution to future real earnings is thus decreasing. In practice, these experimental calculations are quite close to the simple number of enrolled students<sup>12</sup>, and in theory a strongly decreasing value of academic years is not satisfying. Further research, in other countries, would be welcome, including the neutralization of the “creaming off” effect of the best students, which is reflected in the outcome but is not properly speaking an output.

60. The European Credit Transfer and Accumulation System (ECTS) could be another quantity indicator. Built for the Erasmus and then Socrates programmes for facilitating student mobility and international curriculum development, 60 ECTS credits are supposed to measure the workload of a full-time student during one academic year. They are delivered only to “successful” students, achieving more than a certain threshold. They reflect therefore a notion of “quantity” of “(transferred) knowledge and skills”, close also to the notion of “pupils moving up”, exempt of any time-lag or double counting for double degrees. Ideally, a full time curricula of 1 year should not be always valued 60 but 40 or 90 according to the “quality” of the education performed. A notion of individual quality exists in the ECTS grading scale delivered to students, which ranks the students on a statistical basis: A (best 10%), (B next 25%) C (next 30%), D (next 25%), E next (10%) (additional FX and F marks exist for unsuccessful students, FX means: “fail- some more work required to pass” and F: “fail – considerable further work required”), but it is useless for our purpose, as the distribution remains constant over time and consequently any national aggregate.

61. The use of ECTS as quantity indicators would also permit to identify explicitly “imports” and “exports” of education services<sup>13</sup> according to the territory and the nationality of the student (under the assumption he will return to his country after his studies), which would be more convenient for models relying on human capital formation, and informative for the European countries and Commission. Of course its drawback is that it is European only.

62. Our conclusion on higher education is that the concept of credits (like ECTS) would reflect better the education output defined by the expected transfer of knowledge and skills. Our second best could be a weighted number of degrees taking account of their relative “user value”. Our third best is the simple number of enrolled students; it is not quality adjusted, but it is the easiest to calculate, exempt of some drawbacks of the number of degrees, and provides figures very close to the first two bests.

---

<sup>11</sup> Same document (Michael CHRISTIAN, London workshop, October 2006) as above.

<sup>12</sup> 1.3% per year between 1980 and 2001, instead of 1.2% per year for the US, same source.

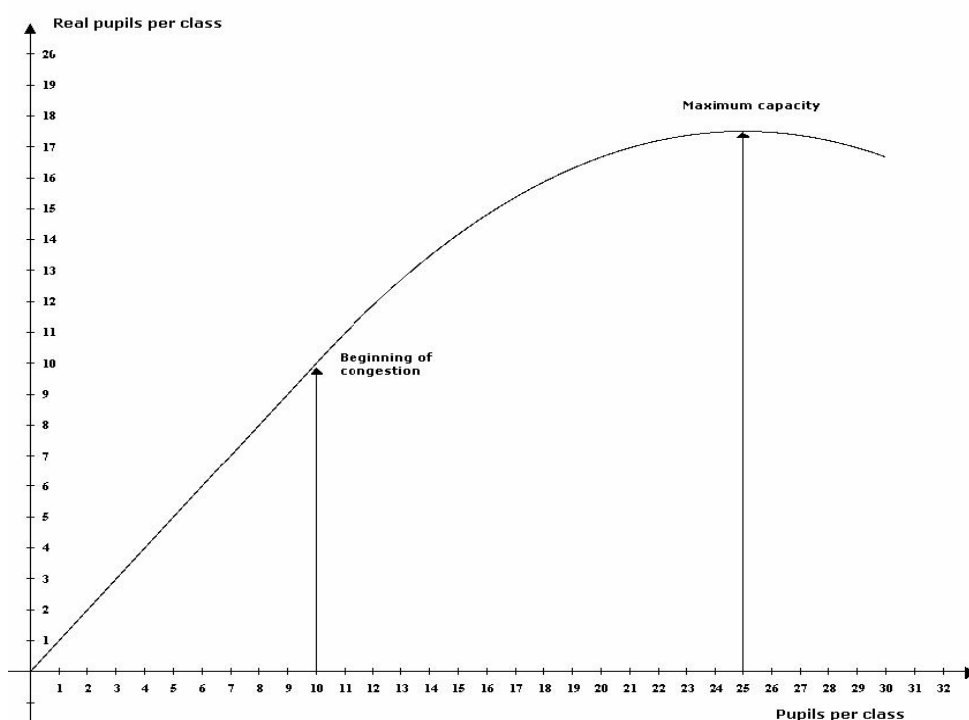
<sup>13</sup> The current SNA records only an actual individual final consumption for the product “education”, and these international exchanges are only implicitly recorded in the global adjustments for “direct purchases abroad by residents” and for “direct purchases in domestic market by non-residents”.

### The use of process or input indicators

63. If outcome-based quantity or quality indicators are not available (scores at exams, degrees, employment rate, etc.), other quality indicators, based on the inputs or on the processes, should be considered. Indeed these factors may have an “expected contribution to the change in the outcome”, and this latter is what has to enter into the calculation of the quality component of the output.

64. Let us take the classical case of the class size<sup>14</sup>, often suggested for a quality adjustment of the education output. The European handbook warns: “*Sometimes pupil/teacher ratios are proposed as indicators of quality. However, it is recommended that this is based on a sound analysis of the relationship between the ratio and the quality of classroom teaching.*” The teacher/pupil ratio is a quality indicator based on the process; if it was applied simply in a multiplicative way, it would provide the same volume index as the classical input method, indeed: number of pupils x number of teachers / pupil = number of teachers!<sup>15</sup> But it can be used in a more clever way, through its expected impact on the outcome. The Italian national accountants (Daniela COLLESI, 2006) have proposed the following model, which assumes a beginning of congestion for the class size of 10 and a complete congestion at the class size of 25: if 25 or more pupils are taught together with the same teacher, it is equivalent as if only 18 pupils at maximum are really taught (7 or more pupils not taught at all), for an equivalent total amount of transferred knowledge and skills for 18 “normal” pupils.

#### **Scheme 2.2.: The “outcome-oriented” Italian model for class-size**



65. With this non-linear pattern, an average national teacher/pupil ratio of 20 is not relevant if numerous classes have a size of 10 and some others of 30. The right use of this quality indicator in a “user value” view is to convert for each class the “number of enrolled pupils” into a “number of real pupils taught”. In fact, the number of “real pupils taught” could be seen as the true “outcome-oriented” quantity indicator, instead of the enrolled pupils (see more details in section 2.3.6, paragraph 161).

<sup>14</sup> The research evidence on the impact of this on student performance is very mixed, which would question its use as the basis for a quality adjustment, but is studied here in a theoretical pattern.

<sup>15</sup> The same if it was used for the stratification with cost weights (see section 2.1.3.1).

66. The Eurostat handbook suggests also a use of school inspections, as they would directly estimate the quality of the output. But such data would need anyway a hard harmonization, then a reduction of the numerous parameters into a single indicator, of which the scaling would be questionable.

67. As discussed in chapter I, the view of this handbook is to recommend using “cost weights” between non substitutable products (consistent with SNA 1993 par. 16.134), while the quantity indicators of substitutable products have to be explicitly quality adjusted (inside the strata) with their “user value”, here the “expected transfer of knowledge and skills” which can be attributed to them.

### 2.1.3.3. Recapitulation

68. The below table recapitulates the recommendations of this handbook for the stratification and the formulas in education output (for the strict education part), combining quantity and quality in each stratum.

**Table 2.7.: Stratification completed with quantity and quality indicators**

Stratum	Quantity	Quality (educational)	Comment
Pre-primary education	Pupil-hours	None	
Primary education: normal	Pupils	Contribution to scores	The sub-stratification normal / special could be replaced by coefficients reflecting the extra costs for social services provided to handicapped pupils
Primary education: special or handicapped pupils			
Lower secondary: normal			
Lower secondary: special or handicapped pupils			
Upper secondary education: general + pre-technical or pre-vocational			Do not forget to subtract the entry education status (can be with the help of a model)
Upper secondary education: vocational			
Post-secondary non-tertiary education			
More practical and occupationally specific programmes tertiary education		Relative future real earnings and employment rate if no scores available	Do not forget the incremental view (contribution of the producer to the outcome). Real earnings and employment rate “without teaching” are to be subtracted.
More theoretically based programmes tertiary education	Credits (ECTS) as 1 <sup>st</sup> best Combination of time-lagged degrees as 2 <sup>nd</sup> best Enrolled students as 3 <sup>rd</sup> best	differentiation by field of education accounting of increasing content of some hard sciences fields relative “value” of B, M, D degrees from labour market?	Different concepts but close figures in practice.

### 2.1.4. Ensuring consistency between market and non-market education estimates

69. Although the first motivation of this handbook was to provide concepts and guidance for individual non-market services, education can also be provided by market producers and the scope of this handbook covers in fact both market and non-market producers. There would be no reason that the measured volume of output measured for education services differs because of the status of the producer when the services produced are exactly similar. Hence, the consistency between the results of the methods used traditionally for the market sector and these used for the non-market output sector has to be examined.

70. The European handbook takes it for granted that market methods relying on Consumer Price Indices or Producer Price indices are the best ones, and the issue of consistency with the non-market methods is apparently not the priority:

*“The education [...] can be delivered as a market or a non-market output. Market output can be treated in a similar way to other market output and therefore much of the following discussion focuses on non-market output.”*

*“[Education] The preferred method for deriving output at constant prices is deflation of value by a price index.”*

*“[Education] Market Services - An A method is deflating output by suitable Producer Price Indices for each type of education service. Appropriate quality adjustments should be made...”*

71. But the breakdown between volume and price is conceptually the same between market and non-market producers if their products are similar<sup>16</sup>. The only practical difference is that in the first case (market) the statistical tradition has been to deflate sales by a price index, and in the second (non market) this chapter recommends a direct volume index, based on cost-weighted quality adjusted quantity indicators. For the two techniques to converge it is essential that the quality adjustments used for deriving the price index for the education market sector are of the same type than the quality adjustments used in deriving the volume indicator for the non market education sector. In other words, the explicit quality adjustments recommended in previous sections for non-market volume measures could as well be recommended for market price measures as generally suggested by SNA 1993 par. 16 118 or by the ILO consumer price index manual par. 7.24-7.28. Classic “matched models” methods without accounting of attainment and achievement, whatever the fine level of stratification, would fail conceptually to measure the right quality change in education.

72. However, it is not granted that the existing price indices used for the market sector of education are indeed based on the same type of quality adjustment. In this context, a case can be made that the best is to use the same method for the market sector than the one used for the non market sector, which is the price-weighted quality adjusted quantity indicators that are proposed in the previous sections. In doing this we would, by construction, verify Principle A of the ATKINSON review: *“the measurement of government non-market output should, as far as possible, follow a procedure parallel to that adopted in the national accounts for market output”*.

---

<sup>16</sup> The framework is perhaps slightly different between market and non-market prices, because the first ones rely on marginal costs and the second ones on average costs, and the demand side is not either regulated by the same individual rationality, but it does not change the basic framework that the volume and price indices should be quality adjusted with the accounting of the user value in each elementary product.

## 2.2. Cross-country comparisons of education output

73. This section is devoted to proposals regarding cross country (or “spatial”) comparisons of the relative levels of the volume of education among OECD countries. It is targeted to be adopted by the PPP exercise among EU and OECD countries, substituting a direct volume estimate to the current price or cost deflation.

### 2.2.1. International databases on education

#### 2.2.1.1. The UOE (UNESCO/OECD/Eurostat) annual data collection

74. The objective of the UNESCO-UIS/OECD/EUROSTAT (UOE) data collection on education statistics is to provide internationally comparable data on key aspects of education systems, specifically on the participation and completion of education programmes, as well as the cost and type of resources dedicated to education (formal education only). These data are detailed by ISCED-97 level and possibly by kind of providers (public education unit, government dependent private education unit, independent private education unit). Altogether, over 60 countries world-wide now complete the UOE electronic questionnaire. Beside this UOE collection, there is a project called WEI (for World Education Indicators) coordinated by the UNESCO-UIS, which concerns 19 countries<sup>17</sup> which could be useful to extend this coverage.

75. The UOE main tables for expenditures are Finance-1: “*Education expenditures by level of education and source and type of transaction*” and Finance-2: “*Education expenditures by level of education, nature and resource category*”. But the concepts of expenditures used in the UOE data collection framework on education finance are not derived from the SNA 1993. The focus is on government expenditures and on financing transactions – taxing, borrowing, spending and lending – rather than on the utilisation of labour, the consumption and the production of goods, and other measures of physical volume. The UOE data collection framework is limited to some elements of current accounts. There is currently a project to examine the differences between the UOE data collection and National Accounts concepts and categories. The UOE data collection provides also several tables on number of enrolled pupils (Enrl-x), number of graduates (Grad-x), number of educational or non-educational staff (Pers-x), number of new entrants (Entr-x) and statistics on average class sizes (Class-x). The UOE data collection already distinguishes between 1<sup>st</sup> degrees (usually bachelors) and 2<sup>nd</sup> degrees (usually masters) in ISCED 5A. Some data are available by fields of education: new entrants, students (including international students) and graduates. Eurostat collects Bachelors and Masters graduates by field of education. These data, for OECD member countries, are usually published in “*Education at a glance*”.

76. A current preoccupation of the UOE collection is the importance of documenting the national sources and methods used in each country in order to provide the data: “*The documentation of data is an integral part of the data collection and is of crucial importance for the future credibility of international education statistics*”, as it is said in the “*OECD Handbook for Internationally Comparative Education Statistics*”, which provides guidance for this UOE data collection.

---

<sup>17</sup> Argentina, Brazil, Chile, China, Egypt, India, Indonesia, Jamaica, Jordan, Malaysia, Paraguay, Peru, Philippines, Russian Federation, Sri Lanka, Thailand, Tunisia, Uruguay, Zimbabwe

77. However, as we can see on the following example, some data are often missing:

**Table 2.8.: Extract of the OECD database on the UOE data collection**

Country	Level of Education	Total expenditure in 2003				FT + PT students in 2003			
		for public institutions	for government-dependent private	for independent private institutions	total	for public institutions	for government-dependent private	for independent private institutions	Total
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
One typical country	1: Pre-primary education	845	793	a	1,638	184,893	214,365	a	399,258
	10: Primary education	2,120	1,972	a	4,092	341,417	411,251	a	752,668
	20: Lower secondary education	xr:30	xr:30	a	xr:30	xr:30	xr:30	a	xr:30
	30: Upper secondary education	2,986	4,179	a	7,165	503,450	716,352	a	1,219,802
	40: Post-secondary non-tertiary education	xr:30	xr:30	a	xr:30	xr:30	xr:30	a	xr:30
	52: Tertiary-type B education	xr(905060)	xr(905060)	a	xr(905060)	xr(905060)	xr(905060)	a	xr(905060)
	905160: Tertiary-type A and advanced research programmes	xr(905060)	xr(905060)	a	xr(905060)	xr(905060)	xr(905060)	a	xr(905060)
	905060: Total tertiary education	1,492	2,085	a	3,577	168,981	209,140	a	378,121
	900000: All levels of education + not allocated by level	7,696	9,132	a	16,828	1,198,740	1,551,108	a	2,749,848
	990000: Not allocated by level	252	104	a	356	n	n	a	n

data extracted on 2006/07/25 09:02 from OECD.Stat

Legend:

a: Not applicable  
m: Missing value  
n: Negligible  
x: Included elsewhere

Many countries have some difficulties to separate ISCED 3 and ISCED 4, or ISCED 5A and ISCED 5B levels of education. However, it should be possible to have most of the countries with decomposition between ISCED 0, ISCED 1, ISCED 2-3-4 and ISCED 5-6 (or ISCED 4 with 5-6 as in ISIC rev 4).

#### 2.2.1.2. PISA (Programme for International Student Assessment)

78. Because our theoretical model to measure the volume output of education is based on the change in the academic scores obtained by students from schools or university, we have a particular interest in international scoring systems, such as PISA and others.

79. The Programme for International Student Assessment (PISA) is an internationally standardised assessment managed by OECD, which is jointly developed by participating countries and administered to 15-year-olds in schools. This age of 15 was chosen as an upper limit of the compulsory scholarship in OECD countries. This three-yearly survey was implemented in 43 countries in the first assessment in 2000, in 41 countries in the second assessment in 2003 and 57 countries have participated in the third assessment in 2006<sup>18</sup>. *“The second survey [PISA 2003] makes comparisons over time possible. This must be approached with caution, however, since two results do not make a trend and since education systems develop relatively slowly.”* Tests are typically administered to between 4 500 and 10 000 students in each country. There are three kinds of examination in PISA: mathematics, reading and science. Each education field is scaled on an average of 500 and a Standard Deviation of 100 among students (hence a relative SD = 20%). The database is rich enough to isolate pupils of public institutions, government-dependent private institutions, independent private ones.

80. An important feature for the use of PISA results in cross country comparisons of the output of education is that a model exists to calculate scores eliminating the specific economic, social and cultural status of the family. This ESCS adjustment will help us to extract from the quality indicator what is not due to the school system.

<sup>18</sup> All OECD countries participate in the PISA exercise, but not all EU members: 8 new member states of EU or candidates did not participate in 2003: Cyprus, Estonia, Lithuania, Malta and Slovenia, as well as Bulgaria, Romania and Croatia. In 2006, only Cyprus and Malta will be missing.

### Box 2.1: The PISA Economic, Social and Cultural Status (ESCS)

The PISA index of economic, social and cultural status (ESCS) is designed to capture broad aspects of a student's family and home background. It is derived from sub-indices based on:

- i) the highest occupational status of the student's parents (ISEI: the International Socio-Economic Index of Occupational Status);
- ii) the highest level of education of the parents converted in years of schooling;
- iii) an index based on home possessions (in fact three sub-indices: the PISA index of family wealth; the PISA index of home educational resources; and the PISA index of possessions related to "classical" culture in the family home).

81. This ESCS correction has been published so far to mathematics only (PISA 2000 and 2003), but similar results can be obtained for reading and science. PISA results are less representative of the quality of the whole (primary and) lower secondary education in some countries, as the limit age for compulsory school is 14 or only the pupils of general upper secondary education are surveyed. In such cases, another correction should be done, relying on the national distributions of PISA scores and on the rate of students following their studies in the curriculum surveyed by PISA, so as to extrapolate what would have been the PISA scores on the whole population of the pupils of former primary and lower secondary schools. More information and figures can be obtained on <http://www.pisa.oecd.org>.

#### 2.2.1.3. Other international student assessments

82. **PIRLS**: the project Progress in International Reading Literacy Study (PIRLS) gives qualitative indicators for primary schools on "reading". 35 countries<sup>19</sup> participated in PIRLS 2001, testing 150 000 9- and-10-year-olds in schools (fourth grade), and 39 countries<sup>20</sup> in PIRLS 2006<sup>21</sup>. This project is managed by the International Study Center (ISC) at Boston College, in partnership with the International Association for the Evaluation of Educational Achievement (IEA) for two major reading purposes: literary and informational. Academic scales are also built on an average of 500 and a Standard Deviation of 100 among all pupils. In the 2006 exercise, PIRLS has provided some "*information on the impact of the home environment on reading achievement and how parents can foster reading literacy.*" A similar correction of the outcome as ESCS for PISA is possible: "*we generally find it more useful to present student scores in mathematics, science, and reading that are unadjusted by social or economic factors. The data are publicly available, and so adjusted scores could be computed*" (Michael O. Martin, TIMSS & PIRLS Co-Director, October 2006). The main drawback for the Eurostat-OECD PPP exercise is the relative small number of OECD and EU countries participating. Another drawback is the heterogeneity of ages between countries, as it is shown in the table below.

---

<sup>19</sup> Of which 22 OECD or EU countries

<sup>20</sup> Of which 24 OECD or EU countries

<sup>21</sup> We can count 42 if United Kingdom (England, Scotland), Belgium (French speaking, Flemish speaking) and Canada (Quebec, Ontario) are counted twice

**Table 2.9.: Average age of pupils by countries for PIRLS assessment  
(sorted by ascending age)**

**TABLEAU I . Âge moyen des élèves de chaque pays**

Average age	Countries
Between 9.7 and 10.1	Cyprus (9.7), Iceland, Belize, Scotland, Italy, Slovenia, Greece, Kuwait, Canada (Ontario, Quebec), Israel, Norway, France, New Zealand, Singapore
Between 10.2 and 10.6	England, Argentina, United States, Hong-Kong, Turkey, Netherlands, Russia, Slovakia, Iran, Germany, Colombia, Czech republic
Between 10.7 and 11.2	Hungary, FYROM, Moldavia, Sweden, Bulgaria, Lithuania, Latvia, Romania, Morocco (11.2)

More information can be obtained on <http://timss.bc.edu/pirls2006/about.html>

83. **TIMSS**: Trends in International Mathematics and Science Study (TIMSS) occurred first in 1995 and then every four years, in the mathematic cognitive domains (i.e. mathematics and science), for 9-year-olds and 13-year-olds (4<sup>th</sup> and 8<sup>th</sup> grades more exactly), involving 76 countries in total but irregularly (France for example only participated in 1995 for the 8<sup>th</sup> grade, so did Germany which will participate again in 2007 but only for the 4<sup>th</sup> grade, and some countries are only surveyed for a fraction: Basque region for Spain, Flemish part of Belgium...) <sup>22</sup>. It is managed by ISC and IEA too. The scale is also built on an average of 500 and a Standard Deviation of 100 among all pupils. Results of 4<sup>th</sup> grade would be complementary to these of PIRLS for primary schools, and the same three education fields as in PISA could be evaluated for primary education if the international coverage was more complete. The estimates for the 8<sup>th</sup> grade could be seen as redundant with PISA, but in the present case they will allow us to extrapolate the missing EU countries in PISA. More information can be obtained on <http://timss.bc.edu/>

84. In the next sections, we propose to use PISA as the base of the main outcome-based quality indicator of the transfer of knowledge and skills for primary and lower secondary education, because of its harmonisation and its extension to all OECD member countries. However, the two other international student assessments can be useful for the other countries involved in the International Comparison Programme.

### 2.2.2. *Computing relative volumes of education services*

#### 2.2.2.1. Model and stratification

85. According to the ideal model we have recommended in section 2.1.3.5., with figures of national accounts in current prices, quantity data (pupils) and weights provided by strata by the UOE collection, and with PISA results as the main quality indicator measuring the average transfer of knowledge and skills (extrapolated if needed by other international students assessments), we suggest a simple application of the following model for spatial comparisons:

$I_{vol}(\text{education unit output}) = \begin{array}{l} I_{quant}(\text{number of enrolled pupils}) \\ * I_{qual ed}(\text{average transfer of knowledge and skills}) \\ + \text{non educational linked services} \end{array} \left. \vphantom{\begin{array}{l} I_{quant} \\ * I_{qual ed} \\ + \text{non educational linked services} \end{array}} \right\} \begin{array}{l} \text{main educational} \\ \text{output} \end{array}$
---

This model is valid for the output and for the Actual Individual Final Consumption.

<sup>22</sup> But Cyprus has taken part to all TIMSS exercises, and Malta will be included in 2007.

**Table 2.10: Model for the spatial comparison of education actual consumption in volume**

Stratum	Quantity	Educational Quality	Non educational linked services
ISCED 0: Pre-primary education	Pupils (pupil-hours?)	None	Should come
ISCED 1: Primary education	Pupils (handicapped counted twice?)	ESCS corrected average PISA score - 180	In a future step from the UOE data collection, based
ISCED 2: Lower secondary: normal	Pupils		
ISCED 3: Upper secondary education	Pupils		
ISCED 4: Post-secondary non-tertiary education	Pupils	None	on the “ancillary services”, after
ISCED 5B: More practical and occupationally specific programmes tertiary education	Pupils	None at present	reconciliation with
ISCED 5A and 6: More theoretically based programmes tertiary education	Pupils (ECTS?)	None at present	NA.

Could be grouped if lack of detail

86. Consistently with our reasoning in section 2.1.3.2., we do not consider any educational quality indicator for pre-primary school. Pupil-hours would be the best quantity indicator, but we suppose that only the number of enrolled pupils is available, as for the other ISCED-97 levels of education in the UOE data collection.

87. We suggest a common educational quality indicator for primary and secondary education based on PISA. The use for upper secondary education is questionable, because PISA surveys pupils at the beginning of this curriculum, but we have thought that it was better to use a quality indicator from previous curricula than no quality indicator at all.

88. We consider that the education outcome is made from 3 contributions: (1) inherited skills, (2) knowledge and skills attributable to the family background, and (3) knowledge and skills transferred by the education units. As our objective is to measure the output of the education services, we have to measure the third contribution separately from the first two. The chart 2.1 shows a typical situation whereby the level of knowledge and skills, measured in terms of PISA scores appears on the vertical axis. As can be seen from the chart, the increase of the education outcome for an individual can originate from inherited skills, family contribution, and the education unit. We would like to measure only the latter, which is the light coloured part of the former chart 2.1 (see section 2.1.2.3., paragraphs 35 and 36).

89. The first factor is the inherited skills. We can assume they are not null since the birth, but they develop while the child grows physically, up to the age of 16 for instance, with more and more neurons and therefore more potential skills, which are not due to school, nor to the family background. The inherited

skills are varying with each child / pupil, they depend on genetics, but we can assume in this exercise they are constant *in average* among countries.

90. The second of these factors is the contribution of the family environment. The analysis of PISA results (OECD 2004, 2005b and 2005c; Hampden-Thompson and Johnston, 2006) confirms the impact of variables measuring students' socio-economic and language background. But the school system is not responsible for it, thus its *output should be evaluated independent from it*. Hence, the PISA raw scores are less relevant for our exercise than the ESCS corrected PISA scores (see section 2.2.1.2. and box 2.1) “as if the family background was the OECD average”. This ESCS correction neutralizes the *varying* part of the family contribution, but not its *worldwide average* value, neither the *average* value of the inherited skills. Our proposed quality indicator for the output therefore becomes  $[ESCS\ corrected\ PISA\ score - X]$ , where X represents the *average* inherited skills plus the *worldwide average* impact of the family background, both at age 15: all what is not due to school at age 15, in average.

91. It is not possible to estimate directly these average contributions within the education outcomes at age 15. But it would be conceptually possible to estimate the sum of them at age 6 on the same academic scale as at age 15, when the school teaching has not yet begun, thus when the education outcome shall reflect what is not due to school at age 6. This education outcome at age 6 is a minimal estimate of our factor X, as both inherited skills and family contribution can go on increasing between ages 6 and 15. We propose then to approximate our quality indicator of the output this way:  $[ESCS\ corrected\ PISA\ score - average\ education\ outcome\ at\ age\ 6]$ .

92. This kind of formula is also consistent with a recommendation of the Eurostat handbook on prices and volume in national accounts: “*it may be appropriate to consider whether examination data (perhaps with a "value added" approach, where pupil abilities are measured before and after a period of schooling) [...] can be used to adjust for changes in quality.*” Conceptually, a pure “exit – entry” indicator will overestimate the value added by the education unit, as we can see on the chart 2.1 (consider the difference between “*value added by the education unit*” and “*exit - entry in the education system primary + secondary educations*”), because of the increasing other factors during the period of scholarship.

93. The worldwide average education outcome at age 6 has been estimated to 180 on the PISA scale, thanks to two national studies: the American National Assessment of Educational Progress (NAEP), conducted since 1969, for mathematics, and the British 4 key stages. They both measure the assessment at different ages on the same academic scale, but they do not test so young people as 6 year old, and they use different conventions for their scaling. Some extrapolations and assumptions have had to be done, as it is explained in the box 2.2 below. The quality indicator is now fixed as:  $[ESCS\ corrected\ PISA\ score - 180]$ .

94. Some may ask whether this constant of 180 is needed as it does not modify the ranking of the countries. But this parameter affects the size of the quality adjustment, which is a ratio:  $[ESCS\ corrected\ PISA\ score\ of\ a\ given\ country - 180] / [ESCS\ corrected\ PISA\ score\ of\ the\ average\ of\ countries - 180]$ . Also, this estimate of 180 may appear on one hand as imprecise and questionable. On the other hand, the direct use of the ESCS corrected PISA score, without any adjustment for inherited skills at age 6, is similar to accepting implicitly that this corresponds to a score of 0 at age 6, which is more questionable. This value of 180 can be seen as a minimum (it neglects the increase of the inherited skills and family contribution during the scholarship), so that  $[ESCS\ corrected\ PISA\ score\ of\ a\ given\ country - 180]$  is necessarily more correct than  $[ESCS\ corrected\ PISA\ score\ of\ a\ given\ country]$ . We have experimented later in this paper two variants, one with 180 and one where this constant is set to zero. We can see that the first one provides a bigger impact to the PISA scores but does not change the picture of the whole output index.

Box 2.2: How to approximately estimate the education outcome at age 6

In theory we would need a survey asking the same questions to pupils of different ages, including up to 6. This does not exist. We have found however two studies that allow deriving a (sketchy) relationship between academic scores and ages: the US NAEP mathematic test which applies on ages 9 and 17 and the UK “four key stages” which gives scores at ages 7, 11, 14 and 16 (in English, mathematics, and science, weighted here equally). We will first assume that these scores can be consistently compared *across ages*. This is of course a very important assumption. The latest results of these surveys appear in columns (1) and (2) of the below table. Our second important assumption is that missing ages can be linearly extrapolated (Scores = a \* Age + b). Using this assumption, we obtain columns (3) for the US NAEP (Scores = 9.5 \* Age + 145.5). As our objective is to benchmark to PISA, which is normalized at 500 at age 15, we normalize proportionally column (3) into column (4). A second transformation is needed because the US test is normalized with a standard deviation of 10% while PISA is normalized with a standard deviation (SD) of 20%. It may look strange that this difference in SD would affect the relative scores *between ages*. However, one must remember that we act here *as if the survey was made at the same moment to all pupils of all ages*. It is therefore logical that the relative average scoring between pupils of different ages is modified by the choice of the standard deviation. In other words, if the pupil of age 6 in column (4), was attributed a score of 352 with a SD of 10%, his/her score under a conventional SD of 20% would be lower, exactly as if this pupil of age 6 was a not very good pupil of age 15. This step therefore consists in modifying the scores of column (4) using a transformation which expands the relative scoring between ages by a factor 2 reflecting the expansion of the SD to 20%. The same operation is conducted on the UK test, without however necessitating this last step, because the UK test is already very close to 1 SD=20%. Overall, the results of the two tests converge on a value of 180 at 6 years old. This is the origin of our proposed constant applicable to PISA scores.

**Table 2.11: Estimate of education outcome at age 6 from the US NAEP and the UK 4 key stages**

<i>Italics = estimated</i>	(1) US NAEP (1 SD = 10%)	(2) UK 4 key stages (1 SD ~ 20%)	(3) Linear US NAEP (1 SD = 10%)	(4) Rescaled US NAEP (age 15=500, 1 SD=10%)	(5) re-scaled US (age 15 = 500, <b>1 SD = 20%</b> )	Linearised re- scaled UK (age 15 = 500, 1 SD = 20%)
Age						
6			203	352	<b>182</b>	<b>180</b>
7		15.51	212	368	217	223
8			222	385	253	267
9	231		231	401	288	310
10			241	418	323	353
11		27.49	250	434	359	396
12			260	451	394	426
13			269	467	429	456
14		33.79	279	484	465	487
15			288	500	500	500
16		35.63	298	516	535	513
17	307		307	533	571	527

95. Using a quality adjustment based on [*ESCS corrected PISA score*] implies that we accept the scaling of the relative knowledge as calculated in PISA. In other words, we accept that if one country has an ESCS PISA score of 600 (the average + 1 SD) and another has a score of 500 (the average), the former has an education system which provided a 20% higher outcome and a 31% higher output if we use the formula  $(600 - 180)/(500 - 180)$ . Note that 20% is consistent with some studies on the economic return of 1 SD in academic scores, through the real earnings (see section 2.1.2.2 and in particular paragraph 17). But some other studies would have suggested a lower economic return, about 8%-10% for 1 SD. If PISA had been normalised with 1 SD = 10%, scores would have all been modified (the distances from 500 would have been divided by two), including the estimate of education outcome at age 6 (340 instead of 180). The new formula would have become [*ESCS corrected PISA score with 1 SD = 10% - 340*], which would have provided exactly the former figures divided by two, and which would have finally led to the same relative quality. Thus the variant of the formula using *differences* in scores over a time period has the

advantage of being independent from the level of the relative  $SD^{23}$ . This does not however suppress all arbitrariness in PISA scales.

96. Of course, if a successful reform of the primary school is conducted in a country, it will result in better PISA scores only *several years* later while the output of the primary school system should have been measured as having increased *since the beginning of the reform*. This is a drawback. Regarding the current estimate, which will be the headline result, the best option is of using the latest PISA estimate, which admittedly will take into account reforms that occurred some years before but not those that have occurred the most recently. With years and continuous new PISA results, the first estimate could be improved several times, up to ten years later! This lagging effect is explained in the below box, and is submitted to experts for a revision policy (for instance “current revisions” during 3 years, then “benchmark revisions” every 5 or 10 years).

Box 2.3: The use of lagged PISA results

PISA measures every three years the outcome of pupils aged 15. There is a time-lag between the measurement of the outcome of a cohort of pupils at age 15, and the output of education provided in primary education (between 9 years ago and 5 years ago) and in secondary education (since 4 years ago). It means that, in theory, the quality of educational services can only be definitively estimated after 10 years! It would suggest successive estimates like in the table 2.12 below, supposing education statuses between two PISA exercises are linearly extrapolated for each grade / cohort and there are the same numbers of pupils across grades. For example: the first estimate for the year 2003 for lower secondary education output would rely on PISA 2003, then, three years later, this estimate would be revised and rely on 1/2 (PISA 2003 + PISA 2006): the first grade of this curriculum would be estimated by PISA 2006, the fourth and last grade by PISA 2003, the second and third grades by an average of both. Note that in any case, first estimates are based on the latest PISA result.

**Table 2.12: Time-lag between outcomes estimates and output quality indicators and revisions**

Years of teaching	Lower secondary education output (hypothesis of 4 grades)	Primary education output (hypothesis of 5 grades)
2000	PISA (2000) then $\frac{1}{2}*(PISA(2000)+PISA(2003))$	PISA 2000 then PISA 2003 then $\frac{1}{5} PISA 2003 + \frac{4}{5} PISA 2006$ at last $\frac{1}{5} PISA 2003 + \frac{3}{5} PISA 2006 + \frac{1}{5} PISA 2009$
2001	PISA (2000) then $\frac{1}{4}* PISA (2000) + \frac{3}{4}* PISA (2003)$ at last $\frac{1}{4}* PISA (2000) + \frac{2}{3}* PISA (2003) + \frac{1}{12}* PISA (2006)$	PISA 2000 then PISA 2003 then $\frac{1}{15} PISA 2003 + \frac{14}{15} PISA 2006$ at last $\frac{1}{15} PISA 2003 + \frac{8}{15} PISA 2006 + \frac{6}{15} PISA 2009$
2002	PISA (2000) then $\frac{1}{12}* PISA (2000) + \frac{11}{12}* PISA (2003)$ at last $\frac{1}{12}* PISA (2000) + \frac{2}{3}* PISA (2003) + \frac{1}{4}* PISA (2006)$	PISA 2000 then PISA 2003 then PISA 2006 then $\frac{6}{15} PISA 2006 + \frac{9}{15} PISA 2009$ at last $\frac{6}{15} PISA 2006 + \frac{8}{15} PISA 2009 + \frac{1}{15} PISA 2012$
2003	PISA (2003) then $\frac{1}{2}*(PISA(2003)+PISA(2006))$	PISA 2003 then PISA 2006 then $\frac{1}{5} PISA 2006 + \frac{4}{5} PISA 2009$ at last $\frac{1}{5} PISA 2006 + \frac{3}{5} PISA 2009 + \frac{1}{5} PISA 2012$

<sup>23</sup> If we have two countries X and Y and PISA scores are noted  $X_{15}$  and  $Y_{15}$ . Then *changes* in PISA scores are  $(X_{15} - X_6)$  and  $(Y_{15} - Y_6)$ . A transformation from one SD to another SD (noted  $SD'$ ) can be mathematically compiled as  $X' = (SD'/SD)* X - Constant$ . This transformation leaves the ratio  $(X_{15} - X_6)/(Y_{15} - Y_6)$  unchanged (i.e.  $(X'_{15} - X'_6)/(Y'_{15} - Y'_6) = (X_{15} - X_6)/(Y_{15} - Y_6)$ ).

#### 2.2.2.2. Number of pupils

97. The UOE data collection should provide directly the number of pupils (full time, part time, full time + part time or full time equivalent) by ISCED-97 levels of education. But, as it has been seen above (see section 2.2.1.1.), some data are missing, either because they rely on public institutions only, either because some ISCED-97 levels are not well distinguished in the national systems, either for other reasons (not applicable, data really missing...), so that some extrapolations will be requested.

98. For pre-primary, post-secondary and perhaps tertiary education, full time equivalent (FTE) number of pupils (or pupil-hours) is more relevant than number of enrolled pupils. For primary and secondary education, our formula is better adjusted with number of enrolled pupils. But a common unit can be preferred if some national figures have to be split between levels of education of the two categories.

99. When the number of pupils is provided for public institutions only, it occurs that the country reports nevertheless expenditures for both “public institutions” and “all institutions”. In this case, the total number of pupils can be extrapolated by the relative expenditures, as if the total costs were the same in public and private institutions.

100. The UOE data collection accepts an item “not ventilated by education level” which is nevertheless belonging to the formal education according to the guidance of the “*OECD Handbook for Internationally Comparative Education Statistics*”. According to the metadata provided by the country, this number can be usually added to the tertiary education or ventilated proportionally.

101. After these usual corrections, some countries have a deficient breakdown of pupils by ISCED-97 level, but have a reliable total number of pupils. In this case the EU or the (EU+OECD) structure can be used. The data are provided in annex 2. 1.

#### 2.2.2.3. Values by ISCED-97 education levels

102. The total amount for formal education has to be taken from national accounts in the PPP exercise, but the ventilation by ISCED-97 levels has to rely on the “total current expenditures” of the UOE data collection, the only international source for such a detailed split<sup>24</sup>. These “total expenditures” have nevertheless to be extrapolated and completed in a lot of cases.

103. Research expenditure has to be subtracted from the more theoretical programmes in tertiary education (5A + 6). The source is provided by a specific table with “core services” and “ancillary services” in “*Education at a glance*”.

104. When countries are restricted to public institutions only for expenditures, it occurs that it reports number of pupils for both “public institutions” and “all institutions”. In this case, the total expenditures have to be extrapolated by the relative numbers of pupils, as if the total costs are the same in public and private institutions.

105. When the expenditures share is deficient, it can be extrapolated either by another year, either by the number of pupils’ structure, at last by the average EU or (EU + OECD) structure.

---

<sup>24</sup> Such a cross country exercise could also be calculated with UOE figures only, but would not be consistent with other figures of the PPPs.

2.2.2.4. A simulation for two countries

106. We will show in this section an example for two countries, Finland and Mexico. A first preliminary remark is that all calculations are made “per capita”: this will allow neutralising the factor of the size of the country (one of the final objectives of this project is indeed to measure GDP *per capita*). The second remark is that all these “cross-country” comparison exercises result in *relative* and not absolute numbers. Thus we will find results such as: Finland = 122.8, Mexico = 100.1. The average level of 100 is relative to a set of countries<sup>25</sup>. However this average is not essential: the important result is that Finland/Mexico = 122.8/100.1 = 1.227. This reads as the level of output of education per capita in Finland is 22.7% higher than in Mexico.

107. *Step 1: Number of pupils and students by ISCED category*

Step 1 consists in calculating a very basic index number simply based on the number of pupils.

**Table 2.13: Number of pupils and students by ISCED category in 1997**  
source: UOE data collection

ISCED level	Finland	Mexico	Average
<i>Total (ISCED 1997)</i>	1 251 049	30 115 758	
<i>0-Pre-primary</i>	181 108	3 432 326	
<i>1-Primary</i>	371 025	14 843 381	
<i>2-Lower Secondary</i>	286 006	6 572 501	
<i>3- Upper Secondary</i>	153 139	3 120 475	
<i>4- Post-secondary non tertiary</i>	4 607	0	
<i>5-6 Tertiary + undistributed</i>	255 164	2 147 075	
<i>Total population</i>	5 194 901	101 396 400	
<i>Total students/population</i>	24.1	29.7	22.11
<b><i>Total students</i></b>	<b>108.9</b>	<b>134.3</b>	<b>100</b>

This results in an index number of Finland of 108.9 (24.1/22.11) and Mexico of 134.3 (29.7/22.11), thus the volume education output per capita of Mexico is equal to +23.3% (134.3/108.9) of Finland. However, this is a very basic measure, which simply reflects the fact that there are relatively more young people in Mexico than in Finland, thus more pupils per inhabitant! It does not in particular take into account the relative *unit cost* of each type of ISCED level.

<sup>25</sup> In this example, this set is technically the EU 25. But this has no impact on the level of Finland versus Mexico

108. Step 2: Stratification

This second step is called “stratification”. The basic idea is that we cannot simply add pupils of all ISCED levels to obtain our quantity indicator. In other words, a student in the tertiary level should not have the same weight as a pupil in the primary level. In accordance with basic volume index number theory, we will weight each pupil by its “price”. The “price” will be taken here as equal to the “unit cost” of each pupil of each ISCED category. The structure of “unit costs” used will be the structure of the average.

**Table 2.14: Shares (in percentage) of expenditures by ISCED level in 1997**

*source UOE data collection*

ISCED level	Finland	Mexico	Average	Average unit cost
<i>Total (ISCED 1997)</i>	100	100	100	1.00
<i>0-Pre-primary</i>	6.6	9.7	9.0	0.71
<i>1-Primary</i>	22.3	36.6	23.2	0.80
<i>2-Lower Secondary</i>	17.6	16.6	19.3	0.84
<i>3- Upper Secondary</i>	23.3	12.7	24.3	1.20
<i>4- Post-secondary non tertiary</i>	0.2	0.0	0.6	0.53
<i>5-6 Tertiary + undistributed</i>	30.0	24.4	23.6	1.69
<b><i>Total students weighted by average unit cost</i></b>	<b>112.2</b>	<b>115.7</b>	<b>100</b>	

The relative volume output of education per capita of Mexico compared to Finland is now +3.1% (115.7/112.2). When relative “price” (or “unit cost”) of ISCED levels are taken into account, countries (such as Finland) that have relatively more students in tertiary education increase in relative level of volume output compared to the previous step. However, this result still does not take into account the quality of the school systems in both countries.

109. Step 3: taking into account the quality of the school system

Here we will use the proposed [*ESCS corrected PISA score – 180*] to obtain the final quality adjusted volume measure. The details of calculation are the following. Starting from the PISA scores of Mexico and Finland in 2000 and 2003, we first apply the ESCS adjustment. We then subtract the parameter of 180 (we have also experimented with the parameter of 0). We then calculate the *lagged* ESCS adjusted scores applicable to the data of each ISCED level 1, 2 and 3 for 1997 (see box 2.3 and table 2.12). We obtain finally the coefficients reflecting the quality of the school system that are applicable to the result of the previous step: for Finland, taking into account PISA results, the coefficient is equal to 1.095, while, for Mexico, this coefficient is 0.866 (both are relative to the coefficient of the average, which is 1, by definition). We obtain the final proposed measure for the relative volume output per capita of the education systems in Finland and Mexico.

**Table 2.15: Taking into account the quality adjustment**

	Finland	Mexico	Average
<i>PISA score 2000</i>	540	410	489
<i>PISA score 2003</i>	545	397	494
<i>ESCS corrected PISA score 2000 (maths)</i>	535	419	493
<i>ESCS corrected PISA score 2003 (maths)</i>	536	419	497
<i>ESCS corrected PISA score 2000 -180</i>	359	264	313
<i>ESCS corrected PISA score 2003-180</i>	357	252	316
<i>Quality adjustment</i>			
<i>ISCED 1: primary : PISA 2003</i>	113.0	79.7	100
<i>ISCED 2: secondary: 1/3 PISA 2000 + 11/3 PISA 2003</i>	113.2	80.1	100
<i>ISCED 3: upper secondary: 8/3 PISA 2000 + 1/3 PISA 2003</i>	114.5	83.9	100
<i>Average quality adjustment applicable to Step 2 result with parameter 180</i>	1.095	0.866	1.0
<i>Average quality adjustment applicable to Step 2 result with parameter 0</i>	1.073	0.911	1.0
<b><i>Final quality adjusted index (parameter 180)</i></b>	<b>122.8</b>	<b>100.1</b>	<b>100</b>
<b><i>Final quality adjusted index (parameter 0)</i></b>	<b>120.5</b>	<b>104.8</b>	<b>100</b>

110. Taking into account quality (with parameter 180), the volume output of Finland per capita is equal to 122.8 compared to 100.1. Thus it is now Finland which is leading Mexico by + 22.7%. Using the implicit parameter of the education level at 6 years as equal to 0, we obtain slightly different results: Finland obtains 120.5 (-2.3) while Mexico obtains 104.8 (+4.7). A calculation was made for other countries with this alternative parameter, showing even smaller impacts on the relative levels.

### 2.2.3. Discussion of this proposal for PPP by international organizations

2.2.3.1. Final report of the Eurostat-OECD Task-Force on Non-Market Services to the joint NA-PPP Working Party in Eurostat (14<sup>th</sup> of May, 2007)

111. The Eurostat-OECD Task Force was in charge of reviewing methods for comparing non-market services volumes and prices in the PPP exercise. In short term perspective, it revised the questionnaire on compensation of government employees. In medium term perspective, it examined an alternative input method relying on national accounts aggregates and ratios, and on ISCO skill levels of qualification, but it was considered that raw data would often be missing or would need an expensive investment. At last, it explored output methods in education and health, and it proposed exactly this method (pupils in FTE from the UOE collection data, PISA quality adjustment for ISCED levels 1, 2 and 3 in the form [ESCS corrected PISA scores – 180]).

112. The final report of the task-force was presented to the joint NA-PPP working party of Eurostat during the 14<sup>th</sup> of May, 2007, and this proposal was welcome by the delegates of the EU member countries, the Eurostat secretariat and the OECD delegate. Few countries expressed some concern for the reliability of PISA. The results were found more plausible than the current input method. An extract of the minutes of this working party can be found in Annex A.

113. The Task Force having completed its mandate, Eurostat and OECD are now responsible for taking this issue further. The analysis will be developed, in particular on the quality adjustment, and several groups consulted, including the OECD group of education experts (23-25 May, 2007), the OECD workshop on measuring education and health volume output (6-7 June 2007), the OECD national accounts working party (October 2007), the Eurostat workshop on volumes and prices (November 2007), the Eurostat national accounts and PPP working parties and the Eurostat Directors of National Accounts Group.

114. An implementation plan is also to elaborate, for the timing of introduction, the length of time series upon introduction, the validation by countries, etc. The new series, potentially available since 2000, will therefore not replace the old ones in the upcoming months.

#### 2.2.3.2. INES Technical Group of OECD education experts (23<sup>rd</sup> – 25<sup>th</sup> of May, 2007)

115. The Technical Group was provided a specific summary of the output method proposal, fully consistent with what was presented in section 2.2.2. The TG:

- WELCOMED the opportunity to review and comment on the proposed methodology
- RECOGNISED the value in moving to an output based method for measuring the output of the education system
- SUPPORTED the proposed use of (UOE) student numbers, stratified and weighted by unit cost as an acceptable quantity measure of education output
- EXPRESSED some concern about the use of PISA scores for the quality adjustment, particularly at ISCED 3 because only limited aspects of output are measured by PISA and ASKED that PISA scores are not used to quality-adjust output at this level
- ACCEPTED with reservations that quality adjustments based on PISA scores are better than no quality adjustment but SUGGESTED that other qualitative measures be investigated including access and progression rates
- ASKED to be further consulted on the calculation and use of education PPPs that would be produced through this methodology

#### 2.2.3.3. The OECD workshop on measuring education and health volume output (6<sup>th</sup> – 7<sup>th</sup> of June, 2007)

116. This workshop had the originality to join experts of several fields: around 30 national accountants, 20 health experts, 10 education experts, another 10 price experts, and at last 10 "transversal" experts, either of productivity or social statistics. The debate focussed more on theoretical wide notions of outcome and on the risk that they could introduce some distortions in the classical framework of national accounts. In practice, the formulas presented for education output and the PPP exercise were little argued, even with the information of the reservations of the INES Technical Group. It was implicitly admitted that the PISA quality adjustment could be restricted to ISCED levels 1 and 2, and that the new output method in education for PPP would have for consequence to transfer the task of collection and extrapolation of data from the PPP experts to the education experts, both in international organizations and in member countries. More details can be read on:

[http://www.oecd.org/document/47/0,2340,en\\_2649\\_34245\\_37733615\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/47/0,2340,en_2649_34245_37733615_1_1_1_1,00.html)

## 2.2.4. Planned developments in international databases and suggestions

### 2.2.4.1. The UOE data collection

117. The concept of “ancillary services” and then “core services” developed by the UOE data collection could improve the comparability between countries, and be useful for national accounts, if the guidelines were adapted from the conventions in national accounts and if the data collection tables were slightly adapted. Very often, the same data set in the Ministry of Education is the common source for national accounts and for the UOE collection, but with very different conventions.

118. The “OECD Handbook for Internationally Comparative Education Statistics” defines in page 62 a classification of the transactions by institutional sectors (in the sense of national accounts) with a focus on “educational institutions”:

*“Goods and Services of educational institutions: all direct public, private and international expenditure whether educational or non educational (e.g. ancillary services) but with one or two exceptions (see below) [not ancillary services => core services];*

*Goods and Services purchased outside educational institutions: private expenditure on educational goods and services;*

*Public subsidies to students for student living costs: regardless of where or how the student spends these subsidies.*

*Excludes R&D outside of educational institutions and private, non-subsidised expenditure on student living costs outside of educational institutions”.*

119. In page 63, this OECD handbook provides the content of the products (in the sense of national accounts) provided by the educational institutions: educational / R&D / non-educational.

#### • Educational Goods and services

- *Instruction (i.e., teaching costs), including in teaching hospitals as it relates to the teaching of medical students;*
- *Educational goods (books, materials, etc) provided by institutions;*
- *Training of apprentices and other participants in combined school and work-based educational programmes at the workplace.*
- *Administration;*
- *Capital expenditure and rent;*
- *Special educational needs; guidance;*

#### • R&D

- *Educational research and curriculum development (including in teaching hospitals- but see below);*
- *Research and development performed at higher education institutions;*

#### • Non-instructional goods and services (Ancillary Services)

- *Student transportation, school meals, student housing, boarding, student health services;*
- *Services for the general public provided by educational institutions;*

120. In our view of national accounts, the definition of the “products” is at least as important as the definition of the “institutional units” and the lack of classification for products bought outside educational units can surprise (but in practice Finance-1 questionnaire asks for some details).

121. Anyway, what is missing in the above definitions vis-à-vis our definition of the scope of education is the parting between “schools” and central government units regulating the educational policy without being directly in contact with the pupils. They are theoretically included in the “educational institutions” in the UOE data collection and they are supposed to provide “educational services” (of which “administration”). On the contrary, they are reported in “general administration” in products in national accounts, not in “education”, or if they have to be, according to the upcoming ISIC rev 4 classification, it will be in another category than “formal education”.

122. Symmetrically, extra school remedial courses, which are not explicitly quoted in the “OECD Handbook for Internationally Comparative Education Statistics” (but which have just been added as “private tutoring” in Finance-1 data collection 2006), are counted as “*Goods and Services purchased outside educational institutions*” in “*Educational Goods and services*”. They are counted in the same industry and in the same product “education” as the regular school system in national accounts.

**Table 2.16: Divergence between national accounts and the UOE collection from a common source**

French codes	In English	For the UOE	In NA (products)
10 enseignement	teaching	Core services	Education
21 cantines et internats facturés à part	Catering, separate invoice	Ancillary	Catering and accommodation
21 b cantines et internats inclus dans facture générale	Catering, included in general invoice	Ancillary	Education (ancillary)
22 médecine scolaire	school health	Ancillary	Education (ancillary)
23 orientation lycées	counselling	Core services	Education (ancillary)
23 b orientation ONISEP	Counselling by a specific public body	Core services	Public administration (or education support)
31 administration générale	general administration	Core services	Public administration (or education support)
32 formation personnel	staff training	Core services	Education
41 transports scolaires	school transport	Ancillary	Transport
42 livres et fournitures fournis par l'école	books and raw material for lessons provided by schools	Core services (educational goods)	Education (ancillary)
42 b livres et fournitures achetés par les élèves	books and raw material for lessons bought by pupils	outside	Books, other goods
43 leçons particulières	remedial courses, private tutoring	outside (private tutoring)	Education
44 habillement et autres fournis par l'école	Clothes and other goods provided by schools	Ancillary	Education (ancillary)
44 b habillement et autres achetés par les élèves	Clothes and other goods bought by pupils	outside	Clothes, other goods

123. “Ancillary services” are used in the UOE collection for both “ancillary services” and “secondary services” in the sense of national accounts. The “ancillary services” should be the non-educational activities provided by the education units, but considered as not separable from the main output, for instance because they have no separate invoices, hence they are recorded in “education” product in national accounts. The “secondary services” are to be reported somewhere else than “education” product in national accounts, for instance because they have separate invoices, and subsequently they are not recorded in “education” product in national accounts. Research and development is a secondary service, and this activity is supposed to be subtracted from education and consistent with the manual of Frascati, both for national accounts and the UOE data collection.

124. In conclusion, “formal education” products and activities in national accounts include extra-school remedial courses (“private tutoring”) and true “ancillary services” but will exclude public general administration and “secondary services”, so that they differ from the concept of “core educational services” in the UOE data collection. The above table 2.16 built with the French “education satellite accounts” reviews the main differences.

125. If the “OECD Handbook for Internationally Comparative Education Statistics” wanted to be fully consistent with national accounts, it could be re-written like that in combining the “institutional sectors” and the “products” classifications:

“• *Goods and Services of educational institutions: in formal education services for ISIC rev 4 (851-853) and NA*

- *Educational services: in education services for NA (ISIC rev 4 groups 851-853)*
- *Ancillary services (not sold): still in education for NA (for instance school health, counselling inside schools)*
- *Research & development: in R&D for NA (ISIC rev 4 division 72);*
- *other secondary services (sold): not in education for NA (catering ...)*

• *General Administration or Education Support: in administration services or education support activities for ISIC rev 4 (841, 855) and NA*

- *Education support activities (ISIC rev 4 group 855)*
- *Ministry of Education (central services) or other specific bodies without pupils but involved in the education policy, not in education for NA (ISIC rev 4 group 841)*

• *Goods and Services purchased outside educational institutions:*

- *Private tutoring: in education services for NA (ISIC rev 4 groups 851-853)*
- *School transportation: in transportation for NA (ISIC rev 4 division 49);*
- *Other educational goods and services: not in education for NA (books, clothes...)*

• *Public subsidies to students for student living costs*

- *not in final consumption for NA”*

If the UOE collection was so detailed, it could then provide the figures for national accounts by ISCED-97 levels of education, and even a non-educational quality adjustment defined either additionally by + *non-educational linked services*, either multiplicatively (second best) by \*  $\frac{\text{total\_xpenditures\_in\_NA}}{\text{core\_educational\_xpenditures}}$ .

#### 2.2.4.2. ECTS and indicators for monitoring the Lisbon objectives in Europe

126. The Lisbon Strategy consists in turning Europe into a modern, dynamic, outward-looking knowledge economy (also called “learning economy”), ensuring that sustainable growth and high quality jobs will follow. Consequently, five education benchmarks were defined for the Lisbon objectives in education and training:

- No more than 10% early school leavers
- Decrease of at least 20% of the share of low-achieving pupils in reading literacy
- At least 85% of young people should have completed upper-secondary education
- Increase of at least 15% in the number of graduates in mathematics, science and technology (MST), with a simultaneous decrease in gender imbalance
- 12.5% of the adult population should participate in lifelong learning

**Table 2.17: List of 29 indicators used by EC between 2003 and 2006 for monitoring education**  
(The indicators for measuring the 5 education benchmarks are in bold)

1) Age of teachers (% of teachers aged over 50 by primary, secondary)	16) Enterprise expenditure on continuing vocational training
2) Number of young people	17) Total expenditure on educational institutions per pupil, in PPS
3) Ratio of pupils to teaching staff	18) Total expenditure on educational institutions per pupil, compared to GDP
<b>4) Completion of upper secondary education</b>	<b>19) Participation in lifelong learning, population 25-64, all, low</b>
<b>5) % of students with low reading literacy proficiency (PISA)</b>	20) Participation in continuing vocational training, all enterprises
6) Performance in reading of 15 year olds (PISA)	21) Participation in continuing vocational training, training enterprises
7) Performance in maths of 15 year olds (PISA)	22) Participation rates in education, students aged 15-24
8) Performance in science of 15 year olds (PISA)	<b>23) Share of early school leavers in population 18-24</b>
9) Participation in education and training of initially low qualified people	24) Distribution of pupils by number of foreign languages learned
10) Students enrolled in MST as a proportion of all students	25) Average number of foreign languages learned per pupil
11) Graduates in MST as a % of all graduates	26) Inward/outward mobility of teachers and trainers, Erasmus+ Leonardo
<b>12) Total number of tertiary MST graduates (growth)</b>	27) Inward/outward mobility of Erasmus students and Leonardo trainees
13) Number of graduates in MST per 1000 inhabitants	28) Foreign tertiary students as a % of all students enrolled, by nationality
14) Public expenditure on education	29) Percentage of the students of the country of origin enrolled abroad
15) Private expenditure on educational institutions	

127. Most of these thresholds objectives are implicitly captured in the quantity indicators retained by level of education (number of pupils / students) or by the educational quality indicator derived from PISA. National accounts match better with quantitative measures (on a linear scale as much as possible) than with binary respect of thresholds, quite arbitrary and suddenly overtaken.

128. The focus on the number of graduates in mathematics, science and technology (MST) suggests an academic view of unequal values of degrees according to the field of education, which could lead to a weighted quantity indicator (either of students, of ECTS or of graduates) reflecting a consequent unequal output, different from unit costs. This idea should be made more explicit by the European Commission or by the academics.

129. Moreover, the European Commission has defined 29 indicators for monitoring the Lisbon Strategy between 2003 and 2006, and more recently 20 new core indicators.

**Table 2.18: 20 new core indicators for monitoring progress towards the Lisbon objectives in education and training**

1) Participation in pre-school education	11) Schools as multi-purpose local learning centres
2) Special needs education	12) Professional development of teachers and trainers
3) Early school leavers	13) Stratification of education and training systems
4) Literacy in reading, mathematics and science	14) Higher education graduates
5) Language skills	15) Cross-national mobility of students in higher education
6) ICT skills	16) Participation of adults in lifelong learning
7) Civic skills	17) Adults' skills
8) Learning to learn skills	18) Educational attainment of the population
9) Upper secondary completion rates of young people	19) Investment in education and training
10) School management	20) Returns to education and training

130. Let us notice the four indicators (26-29) describing the mobility of students (and teachers) in Erasmus programmes and so on. It suggests introducing imports and exports of education in the framework of national accounts, which would be more in line with the human capital theory: the actual final consumption to consider is this of national students (or resident students if they are supposed to work later in the country), wherever they are taught. Hence the actual final consumption should differ from the output. As we have said above (see section 2.1.3.2., paragraphs 60-62), ECTS could provide a first best for an estimate of output in ISCED levels 5A + 6 in European countries and would be the right indicator to manage simultaneously the country of teaching and the nationality of students. International degrees (sometimes counted as double degrees) delivered by a country would lead to double accountings or omissions if they were the direct estimates of “imports” and “exports” (they include a part of their curriculum taught in another country, which would be missed). Of course the simple number of enrolled students is a possible indicator, but we have seen it as a third best. However, ECTS are not explicitly quoted in these 29 indicators.

131. In the very recent list of 20 core indicators for monitoring progress towards the Lisbon objectives in education and training, we can see the “returns to education and training” (last item). It could promote a quality adjustment relying on the “human capital theory”, from the future real earnings on the Labour market (in an incremental view).

#### 2.2.4.3. PISA

132. PISA could expand its scope on other fields of education. The knowledge in a foreign language could be easy to assess, as a variant of “reading” but in another language than the native one. That would allow getting closer to the general balance of scientific and literary matters in secondary education.

133. PISA could also publish the ESCS corrections on all the fields of education it surveys.

134. For certain countries where the surveyed pupils are not representative of children having attended the primary and the lower secondary educations, because age 14 is the end of the compulsory education, PISA could be tested at age 15 and at age 14. A model could provide an equivalent of “age 15 for all pupils having attended the regular school system until the age of 14”.

135. PISA 2009 will introduce the possibility of linking the assessment of 15-year olds with the assessment of students in a younger cohort. The details of this are still being developed and the extent to which countries will take up this option is as yet unknown. It could provide directly a “value added” score for lower secondary education rather than a final score to which we subtract a common large amount (supposed reflecting the education status at age 6). Nevertheless, an intermediate assessment at age 9, 10 or 11 will not modify a lot the quality adjustment of the whole formal education, for it will be positively accounted for the efficiency of primary education and negatively for the efficiency of secondary education.

136. There are no plans to carry out a PISA-style assessment in higher education but the development of a study to assess the competences of adults might in time provide some reference points. Results of such an assessment are some way in the future.

#### 2.2.4.4. Other international student assessments

137. The international coverage of PIRLS and TIMSS, especially 4<sup>th</sup> grade, should develop on EU and OECD countries, in order to envisage a direct quality index for primary education. The average ages of the assessment should be closer (11, for the end of primary education, would be the best).

138. If ESCS-like corrected scores are not directly provided to users, as it is not the intention of the directors of these international tests, the data on the family background are already publicly available and it could suggest to an international organisation calculating these ESCS-like corrected PIRLS and TIMSS scores.

### **2.3. Time series of education output: guidelines and review of good practices**

#### **2.3.1. General guidelines**

139. Much of the guidance has been provided in part 2.1. and especially in section 2.1.3. The following paragraphs summarize the main ideas and give the references in section 2.1. for development.

140. The stratification of the formal education should at minimum distinguish the ISCED-97 levels of education at 1 digit, with a possible merging of levels 5 and 6 (tertiary education). A deeper sub-stratification is suitable in order to reflect the “special needs” of handicapped pupils (if taught in special schools) in first levels of education, the different orientations of the curricula during the secondary education (general / vocational), or the different fields of education / prestige of education units in tertiary education (see section 2.1.3.1., paragraphs 43-52, and table 2.6).

141. The number of enrolled pupils is the basic quantity indicator, for each stratum. If no outcome-based quality adjustment is available, the pupil-hours or the full-time equivalent can be preferred (for instance in pre-primary education, see paragraph 54). If the outcome-based quality adjustment relies on a given grade rather than on a given age, the number of pupils moving up is a more convenient quantity indicator than the number of enrolled pupils (see paragraph 37). Always think about the consistency between the quantity and the quality indicators. In practice, the availability of a quality indicator will determine the suitable quantity indicator with an idea of total knowledge transferred (see paragraphs 30-32). In tertiary education, several quantity indicators can be used: number of credits (first best), number of degrees (second best) or number of enrolled students (third best) (see section 2.1.3.2., paragraphs 56-62 and table 2.7).

142. The educational quality adjustment has to reflect the average transfer of knowledge and skills due to the school system (output = contribution of the economic unit to the change in the outcome) and should ideally be derived from academic scores which are the direct outcome, excluding as much as possible the part due to inherited skills and family contribution. A “value added” measure, where the education status is measured before and after the period of teaching, could provide a good estimate. If no direct estimate of the education status at the entry of the school system is available, the countries could be inspired by the method proposed for the PPPs (see paragraphs 28-29, 33-36, 88-96, chart 2.1, box 2.2 and table 2.11), i.e. the ideal formula should be like: [*final score corrected from socio-economic and educational background* – X], where X represents the average socio-economic and educational background and can be inspired by the values of table 2.11. Of course, time-series have to be guaranteed against any academic drift of the scores.

143. If based on the measurement of the outcome for a given age or for a given grade, the quality adjustment has to take into account the time-lag between the teaching provided and the date of the (final) exam, in a cohort view, which can lead to numerous successive revisions (see paragraphs 36 and 96, box 2.3 and table 2.12), for instance current revisions during three years and benchmark revisions every five or ten years with the other aggregates of national accounts.

144. In the absence of workable scores, other outcome indicators could be used, such as rates of employment and future real earnings, or the delivery of a degree with different values for bachelors, masters and doctorates (see section 2.1.3.2., paragraphs 56-59 and table 2.7), these values being inspired by the indirect outcomes from the labour market if there is no academic consensus on a pure educational scale. As this handbook does not support the human capital theory, the basic model is still that the education system transfers knowledge and skills, and that the individual economic return from the labour market can be assumed proportional to individual knowledge and skills. Under this assumption, the quality of the output is linked with the relative real earnings and not with the absolute ones (so that it will not take benefit from the structural economic growth), in an incremental view (the supplement of future real earnings it provides to students, comparatively to the same young people who would have chosen not to follow studies but to enter directly into the labour market). It is obvious that such economic return can be affected in short term by disequilibrium between supply and demand, so that long-term averages are preferable. It is also clear that relative earnings between graduates and non-graduates reflect for a large part a “creaming effect” which is not an educational output strictly speaking, but which is difficult to estimate. Even if long-term averages are used, the time-lag between the period of teaching and the measurement of the outcome has to be accounted for.

145. Some research on educational outcomes is still needed. For instance, how taking into account the increase of the content of education transferred, due to new knowledge? It should be important for some hard sciences fields in tertiary education. It would be possible to estimate a decrease in comparing two farther years, but much more difficult to estimate a new content, for the past students cannot be assessed retrospectively.

146. If no outcome indicator is available, the inputs or the processes can be taken into account in a quality index, in a form that reflects their expected impact on the change in the outcome (see section 2.1.3.2. and paragraphs 63-65 or further section 2.3.6. and the Italian model for class-size).

147. The non educational goods and services should be counted apart and properly deflated by products, or at least globally estimated in a non educational quality index (second best) (see section 2.1.2.3., paragraph 38 and section 2.2.4.1., paragraphs 117-125).

### **2.3.2 Review of practices in EU-25 and OECD countries**

148. In June 2006, Eurostat and the OECD asked their respective countries to precise the method followed for volume and price in education, and, in case of applied or on process output method, to provide details on the stratification, the quantity and the quality indicators by level of education. No more recent information was collected for this section.

#### *Output methods, including projects*

149. 9 OECD countries had already implemented output measures for education (Australia, Finland, France, Germany, Italy, Netherlands, New Zealand, Spain, United Kingdom) and 7 European countries were supposed to follow by the end of 2006 (Austria, Belgium, Czech Republic, Greece, Hungary, Lithuania, Norway, Slovak republic). Other European countries (excepting Denmark), especially new Member States, expected to apply output methods in the upcoming years. Australia and New Zealand, though non-European countries (which intend to apply the European handbook on volume and price measurement in national accounts), can pretend an anteriority towards most of European countries.

150. In the United States, and also in Korea, these output measures are still a research work, without a precise timing for a concrete application (two American projects were presented in the ONS-OECD-government of Norway workshop in London, 3-5 October 2006). Canada, Japan and Switzerland have no current projects.

#### *Stratification*

151. The ISCED-97 levels of education were the natural reference for stratification, but 6 countries on the 21 concerned had to aggregate some ISCED levels, either pre-primary with primary, either primary with lower secondary, either lower secondary with upper secondary.

152. Some countries use also a sub-stratification of the ISCED levels, for instance 4 countries distinguish the "special" classes for the purpose of disabled pupils, as the unit costs can double (but it is then sometimes difficult to separate primary and lower secondary special education), another 4 countries also split the upper secondary into general / vocational (Finland and France use these two sub-stratifications), and 4 countries also use a regional stratification we do not suggest in principle, except if the curricula are very different and not comparable. 2 countries envisage stratification by gender, because girls and boys are reputed to have different skills in mathematics and reading. This last argument could be approved, in order to better estimate the output within the outcome, but if the gender proportion remains stable over time, it is perhaps a useless sophistication (a regression technique could better take that into account, it could be added to the ESCS components).

153. In tertiary education, 4 countries use a split university / other, 3 countries privilege the field of education (Finland uses a combination of these two dimensions). The institutional sector can also provide another split, but it is not the dimension to privilege.

**Table 2.19: results of Eurostat and OECD questionnaires launched in June 2006**

Countries	Are output methods applied ?	Main sources	Complementary stratification to ISCED 97 ?		Quantity indicator		Weights	Quality adjustment ?
			pre-school -> upper secondary	tertiary	pre-school -> upper secondary	tertiary		
Canada	No current project							
Denmark	derogation until 2012							
Japan	No current project							
Korea	research							
Luxembourg	expected for 2007 ; research in its nascent stage							
Switzerland	No current project							
United States	2 research projects	Mainly surveys	less detail than ISCED : pre-school with primary and lower secondary, perhaps upper secondary too.		number of pupils	number of pupils	costs	several possibilities : pupil-teacher ratio, high school drop-out rate, test scores,
Australia	data since 1989	administrative data	Special ANZICS classification ; pre-school with special, primary with secondary	three HECS (Higher Education Contribution Scheme) bands	number of pupils	number of pupils (+ number of publications for research)	costs	No
Austria	expected by end 2006	administrative data (public) + survey (private)	breaking down for lower secondary between general / vocational + regional stratification	university / other + regional stratification	number of pupils	number of pupils	costs	a combination index of pupils moving up and graduates is used
Belgium	expected by end 2006 (for data from 1995)	administrative data	aggregation pre-school and primary (1), lower secondary and upper secondary (2) + regional stratification	university / other + regional stratification	number of pupils and pupil hours ?	number of pupils and pupil hours ?	costs	No
Czech republic	expected by end 2006	administrative data	Ordinary / disabled ("special") in kindergarten and primary school		number of pupils	number of pupils	costs	No
Finland	data since 2000	administrative data	ordinary, severely disabled, other disabled (primary and lower secondary educations together) + breakdown of upper secondary between general (= <, > 18 year old), vocational (7 fields)	Polytechnic education (7) / university basic degree (20) / higher degree / research	number of pupils + student year for voluntary further vocational training	number of pupils, basic degrees, credit numbers	costs	No

**Table 2.19: results of Eurostat and OECD questionnaires launched in June 2006 (continuation)**

Countries	Are output methods applied ?	Main sources	Complementary stratification to ISCED 97 ?		Quantity indicator		Weights	Quality adjustment ?
			pre-school -> upper secondary	tertiary	pre-school -> upper secondary	tertiary		
France	data since 1981	administrative data + survey	ordinary / disabled + breakdown of upper secondary between general, vocational, apprenticeship	short technical superior (5B) / long superior (5A + 6)	number of pupils	number of pupils	costs	pupils moving up between 1981 and 2001
Germany	data since 1991	administrative data + survey		?	pupil hours (but few updates of hours estimates)	pupil hours (but few updates of hours estimates)	costs	No
Greece	expected by end 2006	survey	aggregation lower and upper secondary		number of pupils	number of pupils	costs ? (specific CPI)	No
Hungary	expected by end 2006	administrative data			number of pupils	number of pupils	costs	class size
Italy	data since 2001	administrative data + survey			number of pupils	number of pupils	costs	class size + degrees / number of students
Latvia	under construction	administrative data + survey	public / private + regional stratification + age and sex ?	public / government dependent private / independent private + regional stratification	number of pupils	number of pupils	number of teachers	class size (primary + lower secondary)
Lithuania	expected by end 2006	survey		stratification by subject	number of pupils	number of pupils	costs	examination data (upper secondary and tertiary)
Malta	after 2006	administrative data + survey	age and sex ?	fields of study	number of pupils	number of pupils		pupils moving up / graduates
Netherlands	data since 1995	administrative data	kind of education (early selection)	NA	number of pupils (corrected for pupils who go to school only partially)	NA	costs	pupils moving up
New Zealand	data since 1978	administrative data + survey	Special ANZICS classification ; but strata less far from ISCED than Australia		number of pupils aggregated with secondary)	number of pupils (full time equivalent)	value added	No
Norway	expected by end 2006	administrative data	owner (CGov, LGov, private) + ordinary / disabled or general (3) / vocational (13) in upper secondary	groups of faculties having different unit costs	pupil hours	number of pupils	costs	on the agenda

**Table 2.19: results of Eurostat and OECD questionnaires launched in June 2006 (end)**

Countries	Are output methods applied ?	Main sources	Complementary stratification to ISCED 97 ?		Quantity indicator		Weights	Quality adjustment ?
			pre-school -> upper secondary	tertiary	pre-school -> upper secondary	tertiary		
Slovakia	expected by end 2006 (later for higher education)	administrative data			number of pupils and pupil hours ?		costs	scoring of pupils ?
Slovenia	hoped for 2008	administrative data + survey	by program		number of pupils and pupils per teacher	number of pupils and pupils per teacher	costs	scoring of pupils ?
Spain	data since 2000	administrative data			number of pupils	number of pupils	costs	class size
Sweden	expected by end 2007	administrative data	boundary between pre-school and family day care ; agregation primary and lower secondary ; stratification ordinary teaching / mothertongue / school meals / school transport / student care	COFOG 0941 : first state of tertiary education and 0942 : second state of tertiary education	pupil hours	COFOG 0941 : number of pupils COFOG 0942 : degree of activity	costs	Number of pupils leaving school without final certificate ; average mark for pupils with final certificate
United Kingdom	data since 1986	administrative data	aggregation of lower and upper secondary ; regional stratification		pupil hours (excluding pupil absences)	number of pupils	costs	estimated contribution that each year of schooling makes to performance in school leaving examinations

### *Quantity indicators*

154. The number of enrolled pupils: 20 countries is more used than pupil hours: 5 countries (some countries use both). But even where pupil-hours are used, they rely on estimates of attainment rarely updated as it is noticed in the German case. Only the United Kingdom can argue it measures directly a true attendance (but it is perhaps not convenient with a quality adjustment relying on scores). 5 countries quality adjust these counts of enrolled pupils by an index of pupils moving up (calculated on a long past period only for France) or getting graduated.

155. 3 countries use number of credits or of degrees in tertiary education (as a quantity indicator or as a quality adjustment)

### *Quality adjustment*

156. On the 21 countries concerned by the output methods, 7 countries admit they do not adjust on quality at all, and have no project for it.

157. 5 countries envisage to use examination data (the US has also listed test scores among its possible quality indicators). 5 countries use or have used an adjustment by the proportion of pupils moving up or being graduated. 4 countries consider the class size as a quality component (Italy has exposed a model, see section 2.3.6, paragraphs 167-170) but we can suspect the other countries use only a ratio teachers/pupils which would lead back to an input method). No country seems to have relied on school inspections.

158. We could have expected more countries using examinations data. There are to our mind two kinds of psychological or material obstacles:

- the fear to confuse output and outcome;
- main data come from national examinations which are regulated by politics. If the purpose is to increase the number of graduates, the “value” of the scores can change over time.

### *Inclusion of non educational goods and services (“ancillary services”)*

159. The OECD non-European countries were asked on the inclusion / exclusion of the cost of all the complementary services like canteens, cafeterias, boarding schools. The general case is that they are implicitly included, as few separate units exist. New Zealand reported some exceptions: a small number of boarding hostels attached to schools and universities. The US said they were included in the General Government sector (S13) but excluded from the Non Profit Institution Serving Households (S15). Only Japan said they were excluded. This aspect is currently not taken into account in the temporal calculations of volume.

### *Weights*

160. All countries have used cost weights, except Latvia, which considers the number of teachers (as a proxy of costs?) and New Zealand, which uses the value added (value added is calculated first by an output method, then the output is obtained by summation of valued added and intermediate consumption, but it seems us incorrect).

### 2.3.3. Could PISA scores be used for national time series?

161. As several countries have thought about change in scores for quality adjustment, but have renounced for psychological or material reasons, an idea is worth exploring: why not using PISA scores for national time series?

162. Each national sample is typically larger than 4 500 but smaller than 10 000 students (some countries like Canada and Mexico use samples of 30 000 students). With such samples, PISA differences (between two countries, but it should be the same between two periods) are considered significant only if they are higher than 7 points. On 27 countries of which the PISA scores can be analysed in 2000 and 2003, 13 countries have an average score changing of more than 7 points in absolute terms. It means that usually we can admit the change between two successive scores is significantly positive or negative (if the difference is lower than 7 points, perhaps it would be better not to take it into account, or only after a checking with another indicator – *a triangulation*). Of course, larger samples would be welcome for a better precision, and we know that some countries take profit of the PISA assessment for testing more students for national purposes. Usually, it is for comparing with younger pupils, or in such a region, etc., but countries could think national time-series too.

163. More precisely, only reading and science have their 2003 scores anchored on the respective reading and science scales from PISA 2000. The published figures in each of the two reports can be directly compared and provide a fair comparison over time for these domains. A consequence is that the mean for reading in 2003 is 494, not 500. For science, it happens to be 500 in both years, but the SD is slightly different.

**Table 2.20: Quality indices and temporal change in quality from PISA 2000 and 2003**

	PISA 2000	PISA 2000	PISA 2003	PISA 2003	Quality change 2000/2003
	reading + science + 2 maths modules	ESCS corrected scores - 180	reading + science + 2 maths modules	ESCS corrected scores - 180	
Australia	526	339	524	336	-0.7%
Austria	510	330	496	315	-4.5%
Belgium	504	326	516	333	2.1%
Canada	527	335	525	334	-0.4%
Czech Republic	500	323	511	329	1.8%
Denmark	497	312	493	305	-2.2%
Finland	538	357	544	356	-0.3%
France	504	332	507	332	0.0%
Germany	486	298	499	321	7.9%
Greece	458	280	463	289	3.3%
Hungary	485	309	491	315	1.9%
Iceland	505	314	498	299	-4.5%
Ireland	509	333	504	327	-1.9%
Italy	471	296	474	299	0.7%
Japan	541	361	530	354	-1.9%
Korea	537	362	541	366	1.0%
Luxembourg	440	269	483	297	10.5%
Mexico	407	261	393	247	-5.3%
Netherlands	..	..	525	346	..
New Zealand	527	336	523	336	-0.3%
Norway	499	304	490	284	-6.7%
Poland	474	309	494	323	4.4%
Portugal	458	298	468	307	3.1%
Slovak Republic	..	..	488	313	..
Spain	485	319	482	312	-2.1%
Sweden	511	319	507	317	-0.6%
Switzerland	505	327	514	337	3.3%
Turkey	..	..	432	298	..
United Kingdom	522	339	..	..	..
United States	492	309	488	295	-4.6%
OECD average	497	317	497	317	0.0%

164. As the mathematic test was deeply renewed in 2000, only two modules can be compared between the two years (space/shape and change/relationships). Results for these two modules in 2000 have been re-calibrated on the 2003 scale.

165. Conceptually, as we have explained in sections 2.1.2.3. and 2.2.2.1., PISA scores have to be ESCS corrected and neutralized of the education status before the entry in the education system (we have estimated as 180 at minimum), in order to better reflect the quality of the output. It is already done by the PISA team for maths only. The above table 2.20, relying on  $[ESCS \text{ corrected scores} - 180]$  formula, provides the set of temporal quality changes we can deduce easily from PISA 2000 and 2003 (before application of the time-lag for the output).

166. Anyway, we can think the best technique would consist in using exhaustive results at exams, if exempt of any drift. The PISA framework (and the US NAEP, and the British 4 key stages quoted in table 2.11) would only provide some guidance: correction of socio-economic and educational factors, subtraction of an “initial” score at the entry in the education system.

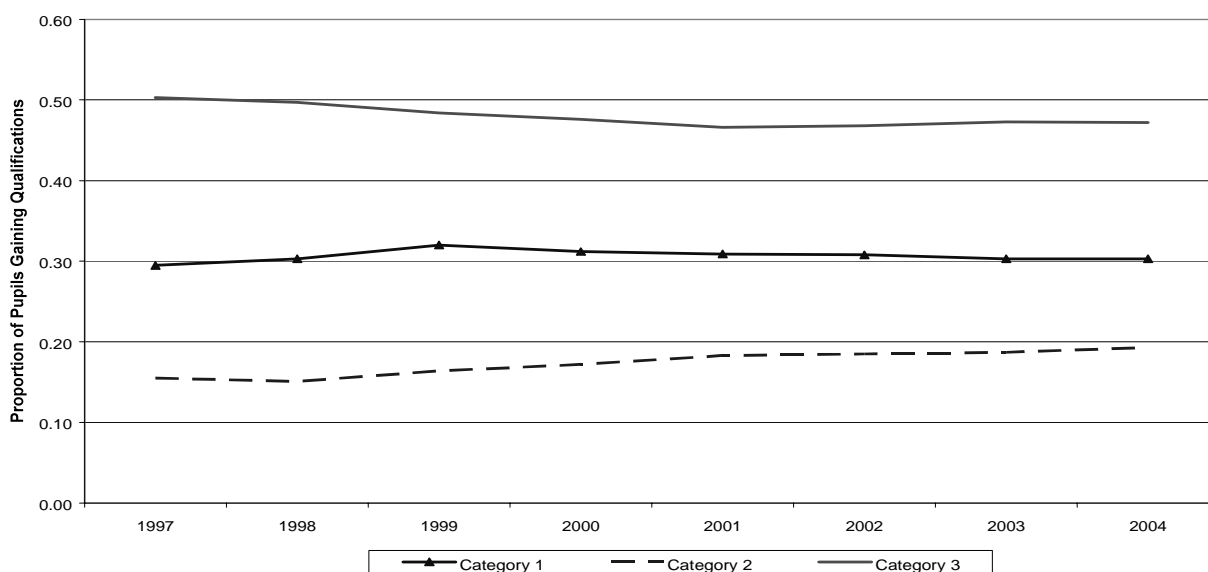
#### 2.3.4. *The Scottish method and a proposal for the use of both examination data and real earnings*

167. Contrarily to the two other national “best practices” that we review in this section 2.3., this method is only a proposal and has not yet been implemented in national accounts, because Scotland is only a part of the United Kingdom, and a consistent framework has first of all to be decided among the “four nations”. All these “best practices” are summaries of more detailed background documents presented during the OECD workshop on measuring education and health volume output (6-7 June 2007), available on: [http://www.oecd.org/document/47/0,2340,en\\_2649\\_34245\\_37733615\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/47/0,2340,en_2649_34245_37733615_1_1_1_1,00.html)

168. This Scottish method is reviewed first because it is the most “outcome-based” one among OECD and European countries, relying on both examination data and future real earnings, which are the most common outcomes explored for the quality adjustment of the output.

169. The basic idea consists in distinguishing 3 categories of students as different “outcomes”/“outputs” of the upper secondary education, including all the qualifications gained over these three years (S4, S5, S6) Category 1 represents the highest levels of qualifications gained; category 2 represents the middle level of attainment; and category 3 represents attainment at the low end of the attainment spectrum:

Chart 2.2 : Proportion of Pupils Gaining Qualifications, by Attainment Category



170. As we can see on the chart 2.2, the main result of these last years is an increase of pupils getting medium attainment and a decrease of pupils getting a low attainment.

171. The second basic idea is to quality adjust these three flows of students with a long-term average of relative future real earnings, taking into account the rates of unemployment:

**Table 2.21: Expected Weekly Wage in Real Terms by Highest Qualification Gained**

	1999	2000	2001	2002	2003	2004	2005	Average Weight
<b>Category 1</b>								
Average Earnings	£275	£283	£304	£298	£290	£297	£308	
Employment Rate	0.75	0.77	0.78	0.76	0.77	0.77	0.79	
<b>Expected Earnings</b>	<b>£206</b>	<b>£218</b>	<b>£238</b>	<b>£226</b>	<b>£223</b>	<b>£229</b>	<b>£242</b>	<b>48</b>
<b>Category 2</b>								
Average Earnings	£211	£209	£230	£202	£229	£206	£212	
Employment Rate	0.72	0.70	0.70	0.71	0.72	0.72	0.70	
<b>Expected Earnings</b>	<b>£151</b>	<b>£145</b>	<b>£162</b>	<b>£143</b>	<b>£164</b>	<b>£149</b>	<b>£149</b>	<b>32</b>
<b>Category 3</b>								
Average Earnings	£176	£196	£203	£201	£188	£209	£210	
Employment Rate	0.47	0.48	0.49	0.50	0.50	0.51	0.50	
<b>Expected Earnings</b>	<b>£83</b>	<b>£93</b>	<b>£99</b>	<b>£100</b>	<b>£94</b>	<b>£107</b>	<b>£104</b>	<b>20</b>

172. This leads to an increasing quality indicator that compensates the demographic decline into an increasing volume output:

**Table 2.22: Quality-Adjusted Output Index for Final Year of Secondary Education**

Number of Pupils	1997	1998	1999	2000	2001	2002	2003	2004
Category 1	18,189	17,899	18,769	18,486	18,583	18,479	18,494	18,315
Category 2	9,557	8,920	9,619	10,191	11,005	11,100	11,414	11,666
Category 3	31,014	29,358	28,388	28,203	28,024	28,079	28,870	28,531
Total Number of Pupils in Cohort	61,659	59,071	58,652	59,250	60,138	59,998	61,035	60,447
Weighted Total	1799210	1731725	1776452	1777500	1804621	1803780	1830318	1823082
Output Index	<b>99.7</b>	<b>96.0</b>	<b>98.4</b>	<b>98.5</b>	<b>100.0</b>	<b>100.0</b>	<b>101.4</b>	<b>101.0</b>

173. The proposal examines then the time-lag issue and wonders if category 1 has to include real earnings got after tertiary education studies.

174. The first point we have to notice there is that the numbers of pupils of each category weighted by the relative real earnings of each category lead to a calculation of the whole real earnings relative to what is obtained by the low attainment. The whole quality adjustment is hence based on the indirect outcome that is future real earnings, in relative terms exempt of the structural economic growth (which is conform to the guidance we have provided in the previous sections), but not in an incremental view. The output provided by upper secondary school should be compared on a real earnings scale, if a direct academic scale is not available, to what would have been obtained without any teaching in upper secondary.

175. Let us imagine that category 3 “low attainment” corresponds in fact to a null output (it is a little excessive but it is said somewhere in the Scottish background document that category 3 means students without qualification). Under this assumption, the coefficients to be affected to categories 3, 2 and 1 are no longer respectively 20, 32 and 48, but 0, 12 and 28, that would provide the same “sign” for the quality adjustment but that would increase its size.

176. The same idea of an incremental view should help to answer the question on the inclusion of students in category 1 following tertiary education and then getting higher future real earnings. Extra future real earnings due to tertiary education have of course to be attributed to the tertiary education only in this framework, and not the upper secondary! The problem is if all students of category 1 follow tertiary education, their future real earnings “without tertiary education” can be unobservable. In this case, a model would be necessary to attribute the part due to secondary education and to tertiary education, including the “creaming effect”.

177. Anyway, the initial question that the Scottish Executive would need to raise is whether the academic scores can be directly observed in the final exams of upper secondary education, for the three categories must rely on scores thresholds. If the scores themselves were used, i.e. the number of points, it would be possible to reason without this indirect indicator that are the future real earnings, needing moreover the use of models to neutralize the influence of further levels of education...

### **2.3.5. *The Finnish case and the use of ECTS credits in tertiary education***

178. The Finnish case relies on a very detailed stratification, especially thanks to the use of fields of education, although not exactly the ISCED ones, which they cross with a distinction university / other. It uses also a split between “normal” (here called “ordinary”) and “special” (here called “disabled” or “special”, two categories) classes in pre-primary, primary and lower secondary educations. In upper secondary education, it distinguishes between general and vocational curricula. These are all good practices (see section 2.1.3.1 and table 2.6).

179. In Finland, two institutional sectors and then two information systems are solicited for the non-market education output: joint municipal boards (S1313) in pre-school, primary, lower secondary, vocational and upper secondary education, polytechnic activity, voluntary cultural activities of adult education centres and similar, basic art education and other educational services (26 indicators) ; state sector (S1311) for universities and some vocational and general education, as well as further education of teachers (34 indicators). The KOTA database provides most central government’s data. The activities of universities account for 95 per cent of central government’s educational service activity.

180. A decision has been made in respect of each level of education as to which indicator(s) is/are the best for measuring its volume of non-market output so that a volume indicator describing the volume of educational services as well as possible can be obtained. The different alternatives for volume indicators have been studied and discussed jointly by experts from the Board of Education and Statistics Finland’s educational statistics. As student-hours were not available, and as the number of degrees was rejected because of the time-lag, the quantity indicators were chosen between number of students, teaching-hours, number of credits and student-years.

181. Number of students was considered as the better established (and even preferred theoretically to student-hours in tertiary education). The number of pupils for a calendar year has been obtained from reported pupil numbers by weighting data on the number of pupils in autumn of the statistical reference year and in autumn of the previous year in the ratio of 50/50. The aforementioned data on pupil numbers become available for the calculations approximately one month after the end of the statistical reference year.

182. Besides number of students, credits are also used for polytechnics and universities. “The specific problem with the use of data on student numbers for tertiary level of education is that students progress at highly varying speeds in their studies and a considerable proportion of them do not necessarily gain a single credit during the academic year. Many of this latter mentioned proportion do not study at all.” Number of credits was then thought considered as better reflecting the educational output and outcome performed by the “true” students. Data on credits for the whole calendar year become available approximately 17 months from the end of the statistical reference year. Data on the spring semester, which become available about five months after the end of the statistical reference year, are used as preliminary data for the latest statistical reference year. We think this use of number of credits is a good practice we recommend as a first best for tertiary education (even if we think that a quality adjustment is still suitable).

183. Student-years are used for the voluntary further vocational training only.

**Table 2.23: Changes in Finnish local government’s educational services**

	Weight in 2003	2001	2002	2003	2004
<b>Pre-school education (within the education system)</b>	<b>0,016</b>	<b>1,246</b>	<b>1,068</b>	<b>0,989</b>	<b>1,012</b>
“Ordinary”	0,674	1,359	1,070	0,995	0,999
Severely disabled	0,070	1,078	1,084	1,023	1,035
Other disabled	0,256	1,011	1,058	0,968	1,040
<b>Comprehensive education</b>	<b>0,734</b>	<b>0,997</b>	<b>1,006</b>	<b>1,005</b>	<b>1,004</b>
“Ordinary”	0,900	0,987	0,996	0,996	0,990
Severely disabled	0,007	0,997	1,046	1,057	1,021
Other disabled	0,030	1,083	1,074	1,065	1,064
Special education	0,063	1,155	1,146	1,112	1,096
<b>Upper secondary general education</b>	<b>0,127</b>	<b>0,990</b>	<b>0,975</b>	<b>0,974</b>	<b>0,979</b>
Under 18-year olds	0,031	0,930	0,926	0,945	0,958
Over 18-year-olds	0,969	0,993	0,978	0,976	0,980
<b>Vocational education</b>	<b>0,050</b>	<b>1,028</b>	<b>1,077</b>	<b>0,995</b>	<b>1,013</b>
Natural resources	0,085	1,020	0,904	1,021	1,063
Technology, communications and transport	0,393	1,021	1,047	1,018	1,002
Social sciences, business and administration	0,096	1,062	1,035	1,009	0,993
Tourism, catering and domestic services	0,125	0,996	1,006	1,010	0,988
Social services and health	0,130	1,028	1,052	1,016	1,025
Culture	0,081	1,123	1,081	1,079	1,060
Recreation and sports	0,004	1,538	1,176	1,129	1,180
Apprenticeship training	0,045	0,961	1,008	1,019	0,964
Further education	0,040		4,644	0,602	1,047
<b>Polytechnic education</b>	<b>0,048</b>	<b>1,078</b>	<b>1,003</b>	<b>1,010</b>	<b>1,003</b>
Natural resources	0,049	1,127	1,181	0,972	
Technology, communications and transport	0,398	1,117	1,063	0,971	
Social sciences, business and administration	0,165	1,027	1,034	0,985	
Tourism, catering and domestic services	0,038	1,332	1,063	1,073	
Social services, health and sports	0,201	1,037	1,010	1,028	
Culture	0,141	1,215	1,229	1,121	
Humanities and education	0,007	1,484	1,153	1,035	
<b>Adult education centres</b>	<b>0,026</b>	<b>1,007</b>	<b>1,007</b>	<b>0,998</b>	<b>1,020</b>
<b>VOLUME INDICATOR</b>	<b>1,00</b>	<b>1,003</b>	<b>1,006</b>	<b>1,000</b>	<b>1,002</b>
Change in volume (old method)		1,012	1,022	1,008	1,019
Difference, volume indicator – old method		-0,009	-0,016	-0,008	-0,017

Number of students

student-years

students

184. Because the Finns privilege the approach by economic unit on the approach by product, the research made by the universities has been kept in the “educational output”, and estimated by the number of publications. This quantity indicator is convenient, but this expenditure should be treated in a separate product, with all the research & development according to the manual of Frascati.

**Table 2.24: Changes in Finnish central government’s educational services**

	Weight in 2003	2001	2002	2003	2004
<b>General education</b>	<b>0.076</b>	<b>0.978</b>	<b>1.010</b>	<b>1.000</b>	<b>1.003</b>
Pre-school education	0.031	0.962	0.947	0.948	0.946
“Ordinary”	0.481	0.962	0.993	1.050	1.028
Disabled education	0.519	1.000	0.922	0.859	0.869
<b>Comprehensive education</b>	<b>0.791</b>	<b>0.975</b>	<b>1.010</b>	<b>1.002</b>	<b>1.003</b>
“Ordinary”	0.640	0.994	0.993	1.004	1.011
Disabled education	0.340	0.937	1.051	0.988	0.988
Special education	0.020	0.924	0.957	1.167	1.028
Upper secondary general education	0.179	1.003	1.023	1.000	1.009
<b>Vocational education</b>	<b>0.021</b>	<b>1.012</b>	<b>1.048</b>	<b>1.088</b>	<b>1.015</b>
Natural resources	0.113	1.057	1.066	1.010	1.056
Technology, communications and transport	0.269	1.000	1.031	1.164	1.036
Social sciences, business and administration	0.102	1.329	1.079	1.000	0.941
Tourism, catering and domestic services	0.250	1.000	1.078	1.109	1.010
Social services and health	0.062	0.611	0.777	1.333	1.042
Culture	0.204	1.044	1.098	1.010	1.000
<b>University education</b>	<b>0.904</b>	<b>1.024</b>	<b>1.045</b>	<b>1.031</b>	<b>1.031</b>
Basic degree programmes	0.383	1.038	1.048	1.027	1.021
Theology	0.009	1.061	1.042	1.114	1.043
Humanities	0.111	1.018	1.015	1.028	1.035
Art and design	0.043	1.056	1.041	1.003	0.942
Music	0.030	0.998	1.051	0.985	0.938
Theatre and dance	0.016	1.055	1.141	1.056	0.938
Education	0.107	1.015	1.017	1.003	1.006
Sport sciences	0.006	1.070	1.111	1.027	1.032
Social sciences	0.068	1.035	1.032	1.020	1.034
Psychology	0.009	1.024	0.995	1.082	0.951
Health sciences	0.011	1.064	1.027	1.076	1.025
Law	0.017	0.990	0.998	1.043	1.021
Economics	0.081	1.057	1.061	1.039	1.024
Natural sciences	0.154	1.058	1.063	1.024	1.025
Agriculture and forestry	0.029	1.005	1.014	1.002	1.031
Engineering	0.200	1.047	1.070	1.020	1.025
Medicine	0.074	1.038	1.062	1.095	1.084
Dentistry	0.008	1.073	1.076	1.080	1.102
Veterinary medicine	0.011	0.995	1.073	1.008	0.908
Pharmacy	0.011	1.076	1.035	0.932	0.951
Fine arts	0.005	1.077	0.955	1.028	1.338
Higher degree programmes	0.252	1.023	1.044	1.047	1.018
Research	0.364	1.011	1.042	1.023	1.089
<b>VOLUME INDICATOR</b>	<b>1.00</b>	<b>1.020</b>	<b>1.042</b>	<b>1.029</b>	<b>1.041</b>
Change in volume (old method)		1.053	1.043	1.015	0.989
Difference, volume indicator – old method		-0.033	-0.001	0.014	0.052

185. This model relies on stable contents of programs and on stable rules for obtainment of credits. We can consider that hard sciences like medicine, economics, technology, etc. always make progress, and this increasing content should be taken into account in the change of volume (further research is needed).

**2.3.6. The Italian case, the use of a class-size model and the accounting of actual time for graduation in tertiary education**

186. Italy has followed the Eurostat handbook on prices and volume in national accounts. As the teaching hours per students are not available in Italy, the quantity indicator is basically the number of enrolled students. In order to build an A method, indicators have been developed to reflect the quality of the service provided in the two major activities.

187. The Italian education output is divided into 4 activities, of which the school system and the university education are the major ones

**Table 2.25: Percentage composition of education output by type of service (current prices)**

Type of service	2000	2001	2002	2003	2004	2005
School system	87.2	86.7	86.5	86.8	85.9	86.2
Vocational training	4.3	4.5	4.3	4.1	4.5	4.4
University education	8.1	8.3	8.7	8.6	9.0	8.8
Subsidiary services to education	0.4	0.5	0.6	0.6	0.5	0.5
Total	100.0	100.0	100.0	100.0	100.0	100.0

188. The school system education is divided into four levels: pre-primary education, primary education, lower secondary and upper secondary education (ISCED-97 levels 0-3). It is supplied mainly in state schools, but other public schools can be managed by local authorities: Municipalities, Provinces and Regions. The volume index is calculated at the maximum breakdown degree. This means that the number of students of state and non-state schools (quantity index) is broken down into the four levels of education and, in upper secondary education, by type of institute (classical lyceum, scientific lyceum, teacher training institutes and schools, vocational institutes, technical institutes, art institutes, art lyceums). This last sub-stratification of upper secondary education is more detailed than recommended in this handbook, but is of the same spirit than suggested for tertiary education, i.e. by field of education (see section 2.1.3.1., paragraphs 40-41, table 2.6).

189. The corrective factor for changes in quality of the service offered by the school system is based on the number of pupils per class. Classes are divided by education level too. The idea is that if the number of students per class increases the individual attention that a teacher may dedicate to each of them decreases. Based on these observations, a conversion function was constructed, which takes on a linear form for the indicator values below the level at which congestion starts ( $q = n$ ), and a non-linear form, similar to a conventional production function, for higher values. Furthermore, the function reaches its maximum in the point of maximum capacity, beyond which the service is overcrowded. The conversion function transforms the actual number of pupils into a number of "standard or real" pupils who have received a service that can be defined as "standard" quality.

$$q = f(n) = \begin{cases} n & 0 \leq n \leq 10 \\ -1/30n^2 + 5/3n - 10/3 & n > 10 \end{cases} \quad \max f(n) = f(25)$$

$n =$  number of pupils

For instance, 25 pupils per class ( $n$ ) allow the teacher to dedicate attention to an equivalent of 17.5 pupils ( $q$ ), which is the maximum reached by this function. The Italian model and the corresponding chart have already been exposed in this handbook and have been said the cleverest ("outcome-oriented") use of the

teachers/pupils ratio, in absence of true “outcome-based” indicators (see section 2.1.3.2., paragraphs 54-56, scheme 2.2).

190. The university education activity has been split in two CPA classes: R&D, deflated apart by an input method, and education *stricto sensu*, stratified in 18 groups of faculties, reflecting the fields of education (both are good practices). A model has been used to determine the unitary cost per student, based on the number of students and on the number of equivalent professors.

191. Italy has chosen two outcome-based indicators, in lack of a unique method proposed by the literature:

- The ratio between the enrolled “regular students” in the course  $SC_{jt}$  (i.e. students who did not exceed the legal length of their degree) and the total number of enrolled students  $S_{jt}$ ;
- The reduction of the distance between the actual number of years for graduation  $LE_{jt}$  and the theoretical length  $LT_{jt}$ .

The corrective factor applied is:  $q_{j,t} / q_j$ , (t-1) where  $q_{jt} = ((SC_{jt} / S_{jt}) + (LT_{jt} / LE_{jt})) / 2$  for each group of faculties  $j$  and each time  $t$ .

The first indicator, close to 1 for “efficient” faculties, reflects the classical “outcome-based” idea of “pupils moving up” rather than “enrolled students”.

The second idea is more adapted to the precise situation of Italian faculties after 1999: the Bologna process had for consequence a re-organization of all the curricula, and in average a reduction of the length of graduation time. It is assumed that the “value” of the final degree has remained constant. At the end of the reform, the ratio should be close to 1, too.

**Table 2.26: Actual and theoretical average time for degree**

Groups of faculties	2000			2004		
	LE	LT	LT / LE	LE	LT	LT / LE
01 sciences	10.18	4.65	0.46	6.90	3.91	0.57
02 pharmacy	11.07	4.98	0.45	7.94	4.75	0.60
03 medicine and surgery	7.52	5.88	0.78	5.88	3.79	0.64
04 engineering	12.78	5.00	0.39	6.97	4.05	0.58
05 architecture	15.05	4.98	0.33	8.02	4.39	0.55
06 agriculture	11.60	5.00	0.43	7.38	4.20	0.57
07 veterinary medicine	12.20	5.00	0.41	8.59	4.75	0.55
08 sociology	7.68	4.15	0.54	7.89	3.80	0.48
09 political science	8.67	4.00	0.46	6.64	3.61	0.54
10 law	8.65	4.00	0.46	8.24	3.81	0.46
11 letter	8.47	4.04	0.48	7.16	3.76	0.53
12 language	8.46	4.00	0.47	6.88	3.72	0.54
13 cultural heritage	8.46	4.00	0.47	8.24	3.83	0.46
14 psychology	10.03	5.00	0.50	6.43	4.24	0.66
15 economics	8.76	4.00	0.46	6.75	3.63	0.54
16 education	7.40	4.09	0.55	6.78	3.82	0.56
17 statistics	7.84	4.00	0.51	6.21	3.45	0.56
18 exercise and sport science	4.00	4.00	1.00	4.48	3.42	0.76
<b>Total</b>	9.75	4.43	0.45	7.00	3.87	0.55

192. The subsidiary services to education are mainly provided by the Bodies for the Right to University Education (ADISU), concerning all the activities supporting the university studies cycle. The main activity, classified here in “education” in the sense of the upcoming ISIC rev 4, is measured in volume by the number of meals supplied and by the number of bed places assigned (“output method”), and the prices are represented at the cost of production of the services. All secondary productions, which are allocated outside of the diagonal in the Supply table, are deflated with a suitable index in relation to the products originating from the secondary productions.

## ANNEX A:

### Extract of the minutes of the joint NA-PPP working party, 14<sup>th</sup> of May, 2007-06-27

#### Final report of the Task Force on non-market services (doc. 07/P1/03)

(...)

The third part of the report – presented by Eurostat as chairman of the Task Force – describes a proposal for an output method for education services. (Health services will be dealt with by a new OECD Task Force.) The Task Force developed an output method both in concept and in practice, by means of test calculations. The method is based on pupil and student numbers by level of education as the quantity measure and a quality adjustment on the basis of PISA, an international student assessment carried out by OECD. The Task Force concluded that the test calculations show that the method provides – on the whole – more plausible results than the current input method. The Task Force proposed to the Working Party to request Eurostat and OECD to develop a plan for the implementation of such an output method in the PPP programme.

The Portuguese delegate asked for further clarification of the choice of a volume approach rather than a price approach, and its relation to quality adjustment. Eurostat said that the approach for education is similar to that used for intertemporal comparisons in national accounts.

The Dutch delegate congratulated the Task Force with the work done and recommended that the results are discussed in a wide audience including education experts. He also suggested to look wider than PISA and to consider for example graduation rates as a source for quality adjustment.

The UK delegate asked whether the results would also be discussed at the forthcoming workshop on quality measurement in November 2007. He asked furthermore to clarify whether the Task Force recommendations were solely for PPP purposes or also for the intertemporal volume measures of national accounts. He said that it should be investigated whether the Task Force recommendations would go against the SNA93 in the sense that output and outcome are mixed up.

The Irish delegate welcomed the new approach proposed by the Task Force and supported a wider consultation of the results of the Task Force, in particular at the November 2007 quality workshop mentioned by the UK delegate.

Eurostat said that the Task Force work will be discussed wider, including at the quality workshop. The proposals however are only intended for the PPP programme and it was felt that PISA is not suitable for use for intertemporal measures. The Task Force took the Handbook on Price and Volume Measures in National Accounts as starting point for the conceptual framework of the method proposed, and that it was therefore thought to be consistent with SNA principles.

The OECD delegate explained that OECD is actively involved in developing methods for non-market services, against the background of a demand for better information on the performance of government. The OECD is preparing a manual of which the first drafts will be discussed in a workshop on 6-7 June in Paris. He emphasised one methodological principle, also adopted by the Task Force, which is that the measurement of market and non-market output should be consistent.

The Finnish delegate said that PISA results are often controversial, even in countries like Finland with a high score in PISA. Nevertheless, Finland would be ready to give Eurostat and OECD the mandate to develop a plan for the implementation of an output method.

The UK delegate asked to inform the meeting about the next steps and how countries will be involved in the decision process. The Dutch delegate asked whether the Task Force will continue working. The Irish delegate asked what a possible implementation plan would contain.

Eurostat and OECD answered by saying that the Task Force has completed its mandate. Eurostat and OECD will now be responsible for taking this issue further. There will be further analysis, in particular on the quality adjustment, and further consultation in the OECD workshop, in the OECD national accounts working party, in the Eurostat national accounts and PPP working parties and in the Eurostat Directors of National Accounts Group.

An implementation plan will – inter alia – have to address questions about the timing of introduction of a new method (e.g. on its own or together with possible other new methods), the length of time series upon introduction, the yearly timetable for the calculations and the way in which countries will validate and approve the results before use in the overall PPP calculations or publication.