Draft minutes of the OECD workshop "Measuring Education and Health volume output"
Paris, château de la Muette, room Roger Ockrent, 6-7 June 2007

1. This workshop was the second one of the OECD "non-market project". The first was in London and organised by OECD, Office for National Statistics and the Government of Norway, 3-5 October 2006. Its main purpose was the discussing of the three draft chapters (chapter 1 on general guidance, terminology and concepts, chapter 2 on education, chapter 3 on health) of an upcoming OECD handbook "Measuring Education and Health volume output" on both temporal and spatial dimensions. About 80 delegates representing 30 delegations were registered, 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. All background documents and presentations are downloadable on the OECD website: http://www.oecd.org/document/47/0,2340,en_2649_34245_37733615_1_1_1_1,00.html. Some delegates were also members of the first meeting of the Health-specific PPP task-force, which took place in the same location on the 8th of June.

Session 1: Measuring education and health output - concepts and guidance (introduction and draft chapter 1 of the OECD handbook)

2. The session was chaired and introduced by François LEQUILLER, who noticed the success of the attendance (80 delegates for 84 seats in room Roger Ockrent). He stressed on the transversal characteristic of the OECD "non-market project", the synergy of the different statistical fields reflected by the various participants, and he pointed out two original features: temporal and spatial dimensions are treated together with the same approach, so are market and non-market producers.

3. Paul SCHREYER presented chapter 1 on general concepts and guidance. In particular, he precised the definition of "outcome" retained in the draft handbook: "level of a state", as most national accountants use it, and not "change in a state" neither "change in a state attributable to the producer", as lot of health experts were used to say, which is closer to the notion of output. He also drew a parallel with the hedonic techniques, sharing a common spirit but here without the help of market prices.

4. André VANOLI discussed chapter 1 and in fact also chapters 2 and 3 for the global consistency of the draft handbook. He regretted that the various meanings of "outcome" were not more discussed in chapter 1. Some economists might include "total utility" or "total economic value" in that meaning, something he would strongly oppose. To extend so much the notion of "outcome-based" measurements of output would distort the purposes of national accounts, by introducing the consumer surplus (which is not in the basic framework of market goods and services). He contested the introductory paragraph of chapter 2 on the links between output, outcome and utility, and commented that they had to be rewritten more cautiously. Nevertheless, he validated the approaches of chapter 2 on education, because "outcome" was taken in the restrictive view of "result of the production process". He addressed some concern with chapter 3 on health services and the potential use of QALYs, attractive concept but out of the National Accounts central framework.

5. The general debate on chapter 1 was quite rich. Francis MALHERBE, of Eurostat, questioned the fact that the use of outcome-based indicators was justified in the OECD handbook by the occurrence of new products, but that the recommendations seem to apply systematically, even when there were no new products.

Isabel QUINTELA, of INE (Portugal), commented the difference between satisfaction, quality and outcome: for instance quality surveys administrated to patients result in focus on waiting time and responsiveness of the medical staff and not quality of the medical service, while they focus on competence when they are administered to medical staff. She would privilege the latter but asks for practical guidance.

Mostafa ASKARI, of Health Canada, asked for explicit identification, in the handbook, of the improvements brought to the current practices of countries. He said also that even if market and non-market producers could be assessed together, it was important to establish different series for different institutional sectors.

Alwyn PRITCHARD, of the UKCeMGA, said that the challenge consisted mainly in defining an objective unit of quality change, and that the policy-makers had to be informed that several measures could suit different purposes: the output for National Accounts, the outcome(s) for other users.

Paul SCHREYER answered that "utility" was used only in explicit quality adjustment.

François LEQUILLER added that the consumer point of view is part of the national accounts, and therefore an accounting for utility is implicitly also part of national accounts. He recognised that, the world of statisticians was divided between "believers" and "non-believers" in these new output indicators, and it was needed to strongly believe in order to achieve some progress.

André VANOLI suggested that output should be distinguished from outcome like inputs are treated differently from output. For the example of drugs helping reducing blood pressure, the output should measure the reduction of blood pressure out of critic thresholds, more than lives probably saved. He used the example of a speed-controlling radar, which should be evaluated based on its technical abilities to measure correctly speed but not on its possible very indirect consequences on the life-expectancy of the drivers.

Isabel QUINTELA said she was perhaps a believer on the concepts, but a non-believer on the practical issues. Health is not a true market where the producers satisfy the consumers' needs, hence utility is a tool which is not adapted. Moreover pharmaceuticals can be outputs for consumers but also inputs for hospitals, hence we need consistent rules.

Alain GALLAIS proposed that a more complete scheme could be developed in Chapter 1 on the gradation between inputs, processes, output, direct outcome, indirect outcome, utility, satisfaction and happiness, in this order. The notion of "outcome" to be used for quality adjustment should be the "direct outcome", the closest as possible to the "result of the production process", and any further concept would introduce problems. The use of utility for quality adjustment was already in the SNA 93 and in the ILO manual on price indices, but the notion of "direct outcome" is the most consistent interpretation of "utility" for the purpose of national accounts.

Bertrand JADOUL, of the Belgian national statistical office, expressed some concern for "the utility of the users of National Accounts": they need simple concepts and formulas.

Paul KONIJN, of Eurostat, declared that he was a believer from the beginning on this work, but that the Eurostat handbook on price and volume did not provide much practical guidance, one of the reasons why the OECD had undertaken this new manual. After the EU regulation, the best practices should be collected among European countries.

Paul SCHREYER confirmed that the concepts and formulas had to be simple and fully understandable.

6. Mark PONT, of the UKCeMGA, exposed the progress in the UK since the Atkinson review. He insisted on the development of satellite accounts in the near future. Isabel QUINTELA asked for documentation and results. Peter SCHERER asked him about the controversy on the health article last year (different proposals with a large bundle of possible evolutions of productivity, positive or negative). Mark PONT answered that the process was still ongoing. A new publication on this topic is planned around September 2007.

Session 2: Education - concepts and international comparisons with an output method in PPPs (draft chapter 2, sections 1 and 2 of the OECD handbook)

7. Alain GALLAIS exposed draft chapter 2 with general concepts and formulas applied to education, based on a definition of "expected transfer of knowledge and skills" itself ideally estimated by the change in scores, essentially in primary and secondary educations. In tertiary education, several good practices are possible, the first best being the number of credits, each credit measuring a unit of "transfer of knowledge and skills". In vocational curricula, the use of future real earnings and rate of unemployment is possible, in an incremental view (contribution of the producer in the change of the outcome). Otherwise, input or process indicators (such as class size) are to be used in an "outcome-oriented" way, after transformation.

8. Paul KONIJN discussed draft chapter 2 with the reference of the Eurostat handbook on price and volume in national accounts, and developed the application of Chapter 2 to the cross-countries comparison, recently adopted for the Eurostat-OECD PPP program. His definition of education is slightly different, relying on a quantity component and a quality component, but it leads to the same formulas. The PPP calculation has been experimented on all European and OECD countries, and the results of the output methods are more reliable than the old results of the input methods. These output methods will therefore be applied in the upcoming years. Nevertheless, the figures are still fragile. Unfortunately, due to confidentiality reasons, the PowerPoint presentation with all its charts cannot yet be uploaded on the OECD website.

9. During the discussion, Francis MALHERBE, of Eurostat, insisted that differences between the OECD Handbook and the existing Eurostat handbook might lead Eurostat not to follow the OECD handbook, because the Eurostat handbook is a European regulation. André VANOLI asked whether the framework for the correction for socio-economic background of pupils were for the whole country or for different units in the same country (as the writing of some definitions of the draft chapter 2 can suggest). François LEQUILLER pointed out that ages 6 and 15 to assess the output in Chapter 2 were conventional, adequate for PISA and PPP, but ideally change in scores was to be assessed every year / grade. Bertrand JADOUL, of Belgium, questioned the frame of the "function" of the change in scores to be counted for in the national accounts. Alain GALLAIS answered first to Francis MALHERBE that the draft OECD handbook definition was consistent in practice with the Eurostat handbook, as Paul KONIJN had said, but was more precise, in this sense that it could avoid inconsistencies between quantity indicator and quality indicator. Second, for the question of Mr VANOLI, the framework and the formulas developed in the handbook were for the whole country and a kind of education services, but they could be applied to different units in microeconomics, and indeed similar formulas were already used to compare the "efficiency" of different units in the same country. Alain GALLAIS said also that curricula were more adapted than single years to measure change in scores, therefore ISCED levels of education were the right object to measure. Coming back to the PowerPoint presentation, he showed that "change in scores" (directly in number of points) was only a proxy of "change in scores due to the education unit", for inherited skills and family contribution could grow during this period and in theory should be excluded.

10. John CRESSWELL, of the Education unit in the OECD, presented PISA in detail, including its lessons on the educational attainment, for instance the big gender difference (girls much better than boys in reading, boys slightly better than girls in maths).

11. Michael DAVIDSON completed his presentation with the comments of the INES technical group of education experts on the new PPP output methods proposed in Chapter 2. Education experts have welcomed the proposals made, even if some were not fully convinced of the quality of PISA. However they preferred that a quality adjustment based on PISA was introduced rather than no quality adjustment. They expressed some concern with the use of PISA for quality adjustment in ISCED level 3 (above all vocational). It would be preferable to limit it to ISCED 1 and 2. Education experts are willing to be more involved in the preparation of the handbook, in the extrapolations for the present gaps in data, as well as for further development.

12. Michael DAVIDSON presented the UOE data collection, with the conceptual differences from National Accounts, and the present gaps in data. Alain GALLAIS completed this presentation with the practical differences between UOE and National Accounts on the scope of "core educational services" for France, and suggested slight adaptations in the UOE aggregates (tertiary approach of output: main / ancillary / secondary rather than main / ancillary, distinction between instructional and non-instructional public bodies…) in order to reconcile UOE and National Accounts aggregates. It could then provide a "non-educational" quality index for the PPP comparisons of education aggregates.

13. Isabel QUINTELA (Portugal) insisted on the fact that education aggregates for the UOE collection had to be made consistent with National Accounts, but not to be integrated in National Accounts. She was supported by Bertrand JADOUL (Belgium) and Daniela COLLESI (Italy), who added that the introduction of ISCED levels of education in National Accounts aggregates had to be made on a voluntary basis. Alain GALLAIS answered that in his view all these calculations have to be developed in a "satellite account", common source for National Accounts and the UOE data collection, providing them consistent and comparable aggregates, but that the central framework of National Accounts was not to be necessarily filled in with all this detail. Alwyn PRITCHARD, of the UKCeMGA, asked how the ISCED levels 2 and 3 (lower and upper secondary education) could be split because of the numerous common education units. Michael DAVIDSON answered that many countries had this problem but were able to solve it, and that they would be asked more and more to extrapolate the missing data by themselves. Paul KONIJN confirmed that in his view UOE and National Accounts aggregates could be made consistent.

Session 3: Education - time series and best practices, other views on education (draft chapter 2, section 3 of the OECD handbook)

14. Richard MURRAY, of the Scottish Executive, presented a proposal for a quality adjusted output method in Scottish education services, relying on future real earnings proportions associated to three categories of pupil attainment of upper secondary education. André VANOLI insisted that the background paper should systematically use the word "knowledge" and then assume that real earnings are linked to knowledge, rather than supposing that real earnings are the outcome of education. Moreover, André VANOLI would be interested by an assessment of education / knowledge with the content which is transmitted and in its link with economic growth, but it is not the view developed in the draft OECD handbook. Alain GALLAIS welcomed the Scottish proposal but questioned the fact that real earnings were taken in absolute terms and not in incremental terms (i.e. only for the part due to the education unit) and that they could reflect tertiary education and not solely secondary education. Real earnings and rate of employment of non-qualified pupils (having not attended upper secondary education) should, ideally, be subtracted. Richard MURRAY pointed out that the weights produced were based on relative rather than absolute differences between the expected future earnings for the 3 categories of attainment. It is not possible to compare the relative difference between the expected

future earnings for these 3 categories of attainment with those who have not attended upper secondary education as all pupils in Scotland will attend upper secondary education (as education is compulsory up to 16 years of age). For the point raised by Mr VANOLI, Alain GALLAIS recognized that such an approach should be developed in hard sciences tertiary education, as similar "scores" at curricula of today and of yesterday have not the same "value" because the content of the transfer of knowledge has changed, but in general this view of contribution to the economic growth was complicated by a lot of external factors and was not consistent with its definition as individual services. Isabel QUINTELA was astonished that very qualified but unemployed students were assumed to have a zero contribution to output. Richard MURRAY clarified that this was not in fact the case as the quality-adjustment is based on attainment. Instead, if a highly qualified student was unemployed then this would impact the weighting given to the relevant category of attainment. Consequently, the student's attainment would still have a positive impact on the quality-adjusted output for education irrespective of whether he or she gained employment.

15. Jani HEIKKINEN presented the Finnish method on tertiary education, relying on a very detailed differentiation of fields of education and the use of credits as quantity indicator. He explained that he was not a "believer" for outcome-based methods, because "outcome" was not a concept of National Accounts and was affected by many external factors. He insisted on a certain paradox that some "soft sciences" fields of education (music, fine arts…) are more costly than "hard sciences" in general, because they needed a higher ratio teachers / pupils and at the same time probably are less prestigious and do not lead to high real earnings. Alain GALLAIS congratulated Statistics Finland for the method on tertiary education, because number of credits can be seen as a direct measurement of output defined as "transfer of knowledge". He suggested also introducing imports and exports of higher education services in the central framework of National Accounts. François LEQUILLER raised the point that even if the maximum differentiation of products seemed a "natural" method, it was insufficient if it assumed that there was no change in quality from year to year, no "new product". So, ideally, a component of quality adjustment is still missing. Aileen SIMKINS, of the UKCeMGA, suggested that fine arts could have a specific social value, and suggested that there was necessity for a quality adjustment based on future real earnings.

16. Daniela COLLESI, of ISTAT (Italy) elaborated on the Italian model for an output method in education, including a function for the class size as a quality adjustment and, for tertiary education, the calculation of unit costs and expectation of degrees by field of education. Paul KONIJN asked whether it was true that overcrowding in classes was a specific Italian issue. Daniela COLLESI confirmed the fact, but said that this issue was progressively solved. Bertrand JADOUL, of Belgium, asked whether it was justified to use an input indicator for quality adjustment in an output method. Moreover, he asked if sometimes more pupils could be translated in a negative productivity. Alain GALLAIS answered that this input indicator was transformed by a model in an "outcome oriented" indicator; hence it was justified in an output method. Yes, some more pupils can be translated in a negative productivity if the Italian model is to be believed, because of overcrowding.

17. Isabel QUINTELA, of INE (Portugal), presented the "education satellite accounts" developed in Portugal, to provide both UOE and National Accounts aggregates. Concepts and rules are those of National Accounts, but some exceptions are introduced: continuous training financed by firms is treated in Actual Final Consumption and not in Intermediate Consumption, imports and exports are identified. Alain GALLAIS asked whether the binary distinction "core educational services / ancillary services" was enough, and if that meant that all ancillary services were treated in the same way in National Accounts, either education aggregates, either other products. The answer was all in education aggregates, but perhaps it could be checked.

18. During the general discussion on draft chapter 2 and time-series, a Dutch delegate asked if ESCS correction had to be applied in time-series also. Alain GALLAIS answered that yes, in principle it should be, it was difficult to measure this change from year to year, but perhaps it was possible to

estimate a trend over ten years, and it could have an important contribution for long term time-series. Aileen SIMKINS said that the stratification between normal classes and special classes was not optimal, as more and more handicapped pupils were enrolled in normal classes. Alain GALLAIS answered that normal classes / special classes was the approach of the ISIC rev 4 classification, but that handicapped pupils in normal classes could be quality adjusted with “ancillary services”. Jani HEIKKINEN, of Statistics Finland, asked whether the output of the central administration (Ministries of education) was part of the OECD handbook. Alain GALLAIS answered they were not in the picture, as they were providing collective services, and a consistent “outcome-based” method would oblige to assess the impact of education reforms, sometimes positive, sometimes negative, of a large magnitude but with poor indicators.

Session 4: Health - concepts and time series in national accounts (draft chapter 3 of the OECD handbook)

19. Sandra HOPKINS summarized draft chapter 3 with the pros and cons for aggregation by illnesses or by institutions, presenting several indicators based on outcome, process or responsiveness.

20. Jack TRIPLETT congratulated her for her contribution. The approach by the cost of disease was sound, and indeed the only feasible one, increasingly accepted, matching scientific data, of course challenging formidable problems. QALY were a little more complex: they were a relevant measure for health outcome, but it did not mean they had necessarily to be the best output indicators. Jack TRIPLETT developed two equations:

    Health = h (medical, lifestyle, environment, genetics) = h (Med, L, E, G)

    Med = f (KLEMS)

    QALY must be in the first equation, but not necessarily in the second which shows the production function of health services where 'Med' is health care output and 'KLEMS' is capital, labour, energy, material and services input. In dentistry, for instance, number of fillings and number of bridges, weighted by their relative prices, provide better indicators than QALY. QALY could provide the same measure as number of treatments, but more imprecisely so, and consequently they should be used only for explicit quality adjustment of new treatments. There are so many new treatments that there is a need for QALY data. But for nursing homes, the number of days in hospital is sufficient. The problems that an output method is challenging are not so different from other industries, for instance co-morbidity belongs to a classical framework of joint costs. Jack TRIPLETT distinguished rules for aggregation of inputs in the first equation, with a marginal social valuation to obtain measures of health / welfare and rules for the aggregation of outputs in the second equation, with marginal costs of production to obtain measures for production in national accounts. He raised also the problem of classifications, which determine the basic units for analysis and should be constructed on analytical principles (not like what happened for the ISIC of which analytical introduction was written in the end!). He considered that COFOG was not the appropriate classification for measuring health output, that ICD-10 was a pertinent classification, and DRGs were more problematic. Above all, Jack TRIPLETT regretted that the imposition of the health functional classification led the US to neglect the cost of illness framework while he recommends developing the latter in priority.

21. During the discussion on the item 4.1, Isabel QUINTELA considered that such frameworks were utopian, because there were no observable data. Alain GALLAIS asked for a clarification / confirmation on the use of QALY: as fillings and bridges are not substitutable purposes, QALY are not justified for a global approach of dentistry, but they would be if different treatments of a same disease had to be provided either in inpatient or outpatient services. Jack TRIPLETT approved. François LEQUILLER congratulated Jack TRIPLETT for his pragmatism and asked him if he could provide

practical examples of using QALYs for specific quality adjustments. Jack TRIPLETT gave a study on heart attacks but was not sure that there existed even a dozen of such studies in the world.

22. Johan POLDER developed a study on cost of illness that he had already presented in the OECD health expert group, on 8 October 2006. He has also achieved a study in Netherlands on the same approach as David CUTLER in the US. Cost of illness framework has to reconcile the top-down and the bottom-up approaches. There are some differences between the Health Providers classification of the Netherlands and that of OECD (for instance, on nursing homes). He presented some results like the difference in costs by gender: females account for 60% of the costs, males for 40%, because of longer life expectancies first, childbirths second. For 5 countries, the results are more comparable at output level than at input level, but it is needed to treat long term care in another framework.

23. Niek KLAZINGA exposed the OECD Health Care Quality Indicators. Some collections are now regular, and provide for instance survival rates. New indicators will focus primarily on patient safety and mental health. When linking quality indicators with volume outputs and cost accounting, cost of illness studies are more relevant than DRGs. He concluded that, while there is undoubtedly progress in health quality indicators, he would not recommend implementing these indicators in national accounts at this stage, because they are still work in progress.

24. Aileen SIMKINS presented the British proposal for a quality adjusted series of general practice (doctors or nurses), by types of consultations, cost-weighted. New statistical method, relying on a sample of practices and no longer on a sample of people, has halved the uncertainty on the activity growth. The article will be published in next September, for discussion.

25. During the general discussion, Charu GARG, of WHO, expressed her concern for the availability of data, especially in developing countries. Why focusing on quality for health services? Pharmaceuticals are already a matter for debate. Jean-Pierre POULLIER asked Johan POLDER in which extent cost of illness studies could be included in national accounts. Johan POLDER answered that COI would be included in the next decade in the Dutch system of health accounts, and that health accounts would be consistent with National Accounts.

Session 5: Round table for conclusions

26. Fernando REIS, of the education unit in Eurostat, was optimistic in the cooperation between health experts and national accounts and said that we had first to draw an exhaustive list of differences between UOE and National Accounts, and that we would see afterwards what could be done. Paul KONIJN insisted on the fact that the OECD handbook should be as practical as possible, and he introduced the upcoming workshop organised by Eurostat on the same topics in November, for time-series in European countries. Eric CHARBONNIER congratulated the organizers and the attendance for the big progress achieved and the good cooperation between statisticians of different fields. For PISA, he expressed some concern and suggested the analysis of access rates and progression rates. For the UOE data collection, the definitions should be checked and clarified. Isabel QUINTELA supported the idea to make the UOE data collection consistent with the National Accounts rules. The Korean delegate asked if price indices in market activities had to be quality adjusted in the same way as these output methods mainly for non-market services. As he was answered yes, he asked for simple and clear methods.

27. Paul SCHREYER concluded the workshop with 5 points:

    - There was general support by participants for the development of volume output measures for health and education. The sense of the meeting was one of pragmatism, however, recognizing that there are many practical and data issues that need resolving before implementation;

    - The mix of participants - national accountants, price and PPP statisticians and health and education experts - turned out to be very helpful in advancing discussions and is probably the only way forward toward tackling complex subject matters such as health and education services;

- The work on the handbook has to live up to the challenge of reconciling the need to follow an evolving area with ongoing research needs and new developments with the request for clear and simple guidance for output measurement;
- The discussions in the workshop - both on education and health - confirmed the significance of classifications and their uses to structure data for meaningful output measures;
- Two main features of the handbook are (a) developing methods for output measurement for producers of health and education services, independently of their status as market or non-market producers; (b) developing methods for temporal comparisons at the same time and consistent with methods for spatial comparisons. This approach towards the development of methods was accepted by workshop participants.

Paul SCHREYER also described the next steps to achieve:

- Revision and expansion of the existing text with a view to presenting a revised draft to national accounts experts (meeting in October), health experts (meeting in October) and education experts (written consultation);
- Close cooperation with Eurostat on this matter (Eurostat organises a workshop in November 2007);
- Organisation of next meeting for Taskforce on health PPPs before the end of 2007.

Participants were given 15 days to send written comments on the three draft chapters.

Alain GALLAIS

21st of June 2007