Paper Prepared for the 31st General Conference of
The International Association for Research in Income and Wealth

St. Gallen, Switzerland, August 22-28, 2010

The OECD Human Capital Project: Progress Report

Marco Mira and Gang Liu

For additional information please contact:

Name: Marco Mira
Affiliation: OECD

Email Address: Marco.MIRA@oecd.org

This paper is posted on the following website: http://www.iariw.org
THE OECD HUMAN CAPITAL PROJECT:
PROGRESS REPORT

Introduction and overview

1. The OECD project on human capital aims to develop an agreed methodology for producing numerical estimates of the stock of human capital for the purposes of international and inter-temporal comparisons. The main features of this project were discussed by CSTAT at its meeting of June 2009 [STD/CSTAT(2009)8]. The methodology used in the project was developed by Dale Jorgensen and Barbara Fraumeni (1989, 1992a, 1992b) and based on discounted lifetime income. The project takes the form of an international consortium consisting of 15 OECD countries (Australia, Canada, Denmark, France, Italy, Japan, Korea, Mexico, Netherlands, New Zealand, Norway, Poland, Spain, the United Kingdom and the United States), two accession countries (Israel and Russia) and one non-member country (Romania). Eurostat and the ILO are also members of the consortium. The consortium is assisted and co-ordinated by the OECD Secretariat. Norway has provided funding for this project.

2. Since 1 October 2009, when the project officially started, the project has proceeded smoothly. In this initial phase, the project has relied on extensive correspondence with consortium partners, to get advice, suggestions, and comments on data and methodology. While the starting point for the project has been various databases already available at the OECD, consortium partners have contributed by providing additional data not available within the OECD.

3. The project, so far, has led to two main outcomes. First, an OECD database has been set up gathering all the essential information needed to measure the stock of human capital for partner countries based on the selected methodology. Second, based on this database, preliminary estimates of the value of human capital for the population of working age have been generated for the majority of the countries participating in the consortium.¹

4. This note provides details on the estimation methodology used; describes the content of the OECD database established for the purpose of this exercise; and highlights some of the preliminary findings emerging from the project. The concluding section describes some of the steps foreseen in the near and more distant future.

Estimation methodology

5. The concept of human capital is a broad one, encompassing a range of attributes (such as knowledge, skills, competencies and health conditions) of individuals. For the purposes of the present project, the term human capital is restricted to people’s knowledge, skills and competencies, i.e. excluding other attributes such as the health status of the population. Further, knowledge, skills and competencies are limited to those acquired through participation in the formal education system, i.e. excluding those gained in the years before primary education and in adult life.

6. Economy-wide measures of the stock of human capital, within the scope defined above, typically rely on physical indicators, such as data on the share of the population with various levels of educational attainment, average years of schooling and, more rarely, test scores of the student or adult population’s competencies. While informative for a number of purposes, these indicators are less suitable for other uses, such as the assessment of the ‘sustainability’ of a development path, which require comparing changes in the aggregate stock of human capital with those in the stocks of other types of assets; such comparisons typically require a common monetary metric.

7. Economic measures of human capital, based on a monetary metric, represent an alternative (or complement) to the physical indicators of human capital. These economic measures can be based on either the value of the inputs that enter the production of human capital (input or cost-based approach) or of the output (typically measured by labour market income) that stems from this investment (output or income-based approach). Following the discussion at the 2009 meeting of CSTAT and the conclusions drawn from literature surveys on human capital measurement (e.g. Liu and Greaker, 2009; Fraumeni, 2008, 2009), the (output or income-based) discounted lifetime income approach was chosen as the preferred methodology to be used in this project for building human capital estimates.

8. The discounted lifetime income approach, advocated by Jorgenson and Fraumeni (1989, 1992a, 1992b), measures the human capital embodied in individuals as the total discounted present value of the expected future incomes that could be generated over the lifetime of the people currently living. Lifetime income is here understood as the labour income associated to different levels of educational attainment.

9. Among other attributes, the discounted lifetime income approach has the potential to be regarded as one important means that may lead to a broader accounting system. In this accounting system, the stock of human capital is measured by lifetime income from both market activities (i.e. the work sold on the labour market) and non-market activities (i.e. household production and leisure time); and the change in this stock in each year (always on the output side) is measured by the change in lifetime income due to changes in the number of people (due to births, deaths and migration), in their education (formal and informal), and in the labour market premiums of education (Fraumeni, 2009).

10. In the current settings, data availability severely limits the possibility of implementing such a comprehensive accounting system. Hence, the present application of the Jorgenson-Fraumeni approach only considers market activities and the labour market returns from formal education. Furthermore, the population is limited to people of working age, i.e. people aged 15 to 64. Within this more narrow scope, the discounted lifetime income approach requires three major steps.

Step one

11. A database containing the economic value of labour market activities for various categories of people needs to be constructed. This database should include, at minimum, information on the number of
people, their earnings (when employed), school enrolment rates, employment rates, and survival rates. All these data should, ideally, be cross-classified by gender, age and levels of educational attainment.\(^2\)

**Step two**

12. An algorithm is needed for calculating the lifetime income for a representative individual in each classified category in the database. The fundamental assumption used here is that an individual of a given age, gender and educational level will have in year \(t+1\) the same labour income and other characteristics (e.g. school enrolment rate, employment rate and survival rate, etc.) as those of a person who, in year \(t\), is one year older but has otherwise the same characteristics (e.g. gender and educational level).

13. The assumption used here distinguishes between three stages in the life cycle of an individual of working age (i.e. between 15 and 64 years of age): i) ‘study-and-work’ (15-40); ii) ‘work-only’ (41-64); and iii) ‘retirement’ (65 and above). Based on this assumption, the lifetime labour income of an individual can be computed as follows:

- For persons aged 65 and over (i.e. ‘retirement’ stage), their lifetime labour income is zero since, by assumption, these persons will not receive earnings after withdrawing from the labour market.

- For persons aged 41 to 64 (i.e. ‘work-only’ stage), their lifetime labour income is estimated as follows:

\[
LIN_{age}^{edu} = EMR_{age}^{edu} AIN_{age}^{edu} + SUR_{age}^{edu} + LIN_{age+1}^{edu} \left[(1 + r)/(1 + \delta)\right],
\]

where \(LIN_{age}^{edu}\) is the present value of lifetime labour income for a representative individual with educational level of “edu” at the age of “age”; \(EMR_{age}^{edu}\) is the employment rate for this individual; \(AIN_{age}^{edu}\) is his/her current annual labour income; \(SUR_{age}^{edu}\) is the probability of surviving one more year given that this individual is at the age of “age”; \(r\) is the annual growth rate of the labour income (in real terms) of a person of these characteristics in the future; \(\delta\) is the annual discount rate.

The lifetime income of a representative individual during the ‘work-only’ stage is therefore estimated as the sum of two parts: the first part is the current labour income, adjusted by employment rate (the first term in equation (1)); the second part is the lifetime income in the next year, adjusted by the corresponding survival rate, income growth rate and discount rate (the second term in equation (1)).

- For persons aged 15 to 40 (i.e. ‘study-and-work’ stage), their lifetime labour income is estimated as follows:

\(^2\) In practice, most data on survival rates do not distinguish between different categories of educational attainment (i.e. survival rates differ only according to the age and gender of each person).
\[
\text{LIN}_{age}^{edu} = EMR_{age}^{edu} AIN_{age}^{edu} + \left(1 - \sum_{edu} \text{ENR}_{age}^{edu} \text{t}_{edu} \right) \text{SUR}_{age+1}^{edu} \text{LIN}_{age+1}^{edu} \left[\left(1 + r\right)/(1 + \delta)\right]
\]

\[
+ \sum_{edu} \text{ENR}_{age}^{edu} \text{t}_{edu} \left[\sum_{i=1}^{\infty} \text{SUR}_{age+i}^{edu} \text{LIN}_{age+i}^{edu} \left[\left(1 + r\right)/(1 + \delta)\right]^i \right] / t_{edu - edu},
\]

where \(\text{ENR}_{age}^{edu} \) is the school enrolment rate for a representative individual with educational level of “\(edu\)” pursuing his/her studies into a higher educational level of “\(edu\)”; \(t_{edu - edu}\) is the school duration for this individual with educational level of “\(edu\)” to complete a higher educational level of “\(edu\)”. 

During the ‘study-and-work’ stage, a representative individual in the next year will be confronted to two courses of action: the first is to continue his/her work (holding the same educational level as before) and earn income as \(\text{SUR}_{age+1}^{edu} \text{LIN}_{age+1}^{edu} \left[\left(1 + r\right)/(1 + \delta)\right]\), with the probability of \(\left(1 - \sum_{edu} \text{ENR}_{age}^{edu} \text{t}_{edu} \right)\); the second is to enter into school and (after completing study having gained a higher educational level) to receive income as \(\left[\sum_{i=1}^{\infty} \text{SUR}_{age+i}^{edu} \text{LIN}_{age+i}^{edu} \left[\left(1 + r\right)/(1 + \delta)\right]^i \right] / t_{edu - edu}\), with the probability of \(\sum_{edu} \text{ENR}_{age}^{edu} \text{t}_{edu}\).

Therefore, his/her lifetime income in the next year is the expected value of the outcomes of these two courses of action (i.e. the sum of the second and the third terms in equation (2)).

14. The empirical implementation of equations (1) and (2) is based on backwards recursion. In this approach, the lifetime labour income of a person aged 64 (i.e. one year before retirement) is simply his/her current labour income (the first term in equations (1) and (2)) because his/her lifetime labour income at 65 is zero by construction. Similarly, the lifetime labour income of a person aged 63 is equal to his current labour income plus the present value of the lifetime labour income of a person aged 64, and so forth.

15. In estimating lifetime labour income by using equations (1) and (2), several practical assumptions are made, some of which are used as well by other studies in the field (e.g. Gu and Wong, 2008; Le et al., 2006; Liu and Greakeer, 2009; Wei, 2004, 2007). The most important assumptions are as follows:

- Individuals can only enrol in a higher educational level than the one they have already completed.
- No further enrolment is allowed for people having already achieved the highest educational level.
- Students enrolled in educational institutions requiring more than one year to complete are assumed to be evenly distributed across the total study-period. This is equivalent to say that, during each school-year, there is the same (equal) proportion of the total students that will complete the study.
- No delaying and quitting are allowed during the whole study period.
**Step three**

16. The lifetime labour income measures estimated through equations (1) and (2) are applied to all individuals in each age/educational categories to compute the human capital stock for each category. Summing up the stocks of human capital across all classified categories yields the estimate of the aggregate value of the human capital stock ($HC$) for each country.

$$HC = \sum_{age} \sum_{edu} LIN_{age,edu} N_{age,edu},$$

where $N_{age,edu}$ is the number of persons in the corresponding age/educational category. Note that equations (1), (2) and (3) are applied separately to both males and females, which allows computing the stock of human capital by gender.

**Database construction**

17. Although the discounted lifetime income approach ideally requires information by single year of age, all data available within the OECD system (as well as most data available to researchers in individual countries) refer to (either 5-years or 10-years) age groups. To develop a database suitable for this exercise, the OECD Secretariat has relied on a number of practical assumptions and imputation methods to generate data by single year of age; for example, data on the number of students by single year of age were based on information on the average enrolment rate of a given age group, and on the number of people of each age. These assumptions and imputations obviously affect the quality of the estimates shown here.

18. A further factor shaping estimates of the stock of human capital is the quality of the underlying data in various years. In general, the quality of the OECD Education database is lower for the years preceding the early 1990s. In particular:

- School enrolment data for most countries are of better quality starting from 1998, partly reflecting changes in the International Standard Classification of Education (from ISCED67 to ISCED97) around that year.
- Data on annual earnings by educational attainment categories are only available since 1997 for most countries participating in the project.
- Similarly, the OECD Education database contains two types of educational attainment datasets: the first, with more detailed educational categories, starts for most countries from the 1990s; and the second, with more aggregated categories, provides time-series going back longer in time.

Because of these factors, it was decided that, in the first stage of the project, estimates of human capital would be limited to starting from the 1990s (1997 for most countries), where data are of better quality.

19. For each country participating in the consortium, a database was established consisting of the following five data files covering the various elements that enter the estimation of the human capital stock based on the chosen methodology.
Survival rates

20. Information on survival rates, by gender and individual year of age, was mainly derived from country life-tables published in the Human Mortality Database (http://www.mortality.org). For a few countries that are not included in this database, mortality data were obtained through bilateral contacts with the country representatives.

Educational attainment

21. Information on the number of people by the highest educational level attained is available through the OECD Education database managed by the OECD Education Directorate. These data are mainly based on national labour force surveys (LFS). For most of the countries, these data are available by gender, educational level and 5-year age groups for people within age 15 to 64.

22. Following consultation with members of the consortium, it was decided to focus on this working age population (age 15 to 64) for the purpose of the project at the current stage. Two adjustments were applied. First, data on the working age population by educational attainment were benchmarked to the levels (by gender and 5-years age groups) of the total population available in the OECD demographic database. Second, to break down these data referring to 5-year age groups into data by individual year of age, national data on the population by gender and individual year of age were used to interpolate across different educational categories.

Employment rates

23. OECD data on employment rates by gender and educational level are based on the same source (Labour Force Surveys) used for educational attainment. Data on employment rates by 5-year age group were broken down in data by individual year of age based on the assumption that the employment rates for each 5-year age group apply to each single year within the corresponding age group.

School enrolment rates

24. Information on the number of students by age (by single year of age up to 29, by 5-year for age groups 30-34 and 35-39, and by one group for 40 and above) enrolled in different educational institutions classified according to ISCED97 can be directly downloaded from the OECD online databank (http://dotstat.oecd.org/wbos/Index.aspx). Combined with educational attainment data, these data on the number of students allow computing school enrolment rates, defined as the share of people who completed a given level of education and enrolled in the level above. For simplicity, the school enrolment rates were computed only up to age 40. For people aged 30 and above, the assumption made to obtain data by individual years of age is that the school enrolment rates for each single year within age groups, 30-34, 35-39 as well as 40 and above are the same as those for the corresponding age groups.

Annual earnings

25. Data on annual earnings by gender, age groups and educational attainment of workers are available through the OECD Education database. The original sources of these data vary across countries (labour force surveys for some countries, household income survey and others population surveys for other countries). Partially for this reason, the reference period for earnings differs across countries. While some countries provide estimates of annual earnings (based on weekly and monthly data included in the original

---

3. The survival rate is the conditional probability that a person who is alive in year \( t \) will also be living in year \( t+1 \).
26. The earnings data in the OECD Education database may also reflect differences in how part-time and full-time workers (as well as students) were treated by the national correspondents providing estimates to the OECD. A more detailed assessment of these country differences in earnings’ definitions (e.g. whether they include employers’ social security contributions) will be pursued in the next few weeks.

27. Because of uncertainties on the exact definition of earnings reported in the OECD Education database, earnings data by educational attainment were benchmarked on the series ‘wages and salaries’ from the OECD annual national accounts. This implies that data on the ratios between the earnings for different educational categories (from the OECD Education database) were applied to the national account series of wages and salaries per employee; this procedure has allowed deriving estimates of the value of annual earnings by gender, age groups and educational levels consistent with SNA totals.

28. Two additional adjustments were applied to the OECD earnings data now benchmarked on SNA totals. First, in order to obtain earnings by single year of age, a parabolic curve was fitted through the observations of annual earnings by age groups (15-24, 25-29, 30-34, 35-44, 45-54, 55-64). Second, as the educational categories in the OECD annual earnings datasets do not always match those in the educational attainments datasets, imputations were used to generate annual earnings anchored to the educational levels shown in the educational attainment data sets.

Preliminary results

29. Preliminary estimates of the stock of human capital have, so far, been computed for thirteen OECD countries (Australia, Canada, Denmark, France, Italy, Korea, Netherlands, New Zealand, Norway, Poland, Spain, the United Kingdom and the United States), one accession country (Israel) and one non-member country (Romania), across several years. For simplicity, all estimates are based on common assumptions on the growth rate of real earnings of each age/educational categories in the future (1.32% per year); and of the discount rate (4.58%). These two values are used by Jorgenson and Fraumeni (1992a) in their estimations for the United States.\footnote{Several studies have shown that human capital estimates are sensitive to the choice of these parameters (e.g. Wei, 2004; Gu and Wong, 2008, 2009; Liu and Greaker, 2009). The purpose of choosing 1.32% and 4.58% as used by Jorgenson and Fraumeni (1992a) in our experimental calculations is to show the potential of comparative analysis. More explorations on the choice of these two key parameters should be carried out in the future.}

30. A first set of results relates to earnings profiles by age and educational attainment. These have been computed separately for males and females. Figure 1 and 2 present estimates for Canada in 2006, based on the two concepts of annual labour income and lifetime labour income. The age-earnings profiles for some other countries are shown in Appendix. These figures highlight a number of patterns that were also identified in a range of national studies (e.g. Le et al., 2006; Liu and Greaker, 2009; Wei, 2004, 2007). In particular:
First, both annual and lifetime labour incomes rise with age and then gradually decline for all educational levels. The peaks occur at younger ages for lifetime income than for annual income.

Second, both annual and lifetime labour incomes are higher for higher levels of educational attainment.

Third, both annual and lifetime labour incomes are higher for males than for females, at all levels of educational attainment.

Comparisons of the value of human capital per capita in 2006 (in US$ at PPP rates) are shown in the left panel of Figure 3. The estimates shown here are for illustrative purposes only and they reflect the (unrealistic) assumption that the real growth rates of annual earnings and the annual discount rate are the
same in all countries. Based on these (provisional) assumptions, our estimates indicate that values of human capital per capita range between 78.65 USD in Romania and 642.69 USD in the United States.

Figure 3. Human capital per capita and ratios of total stock of human capital to GDP, 2006

Note: Estimates for Australia refer to 2001 and for Denmark to 2002.

32. An alternative yardstick for comparing the stock of human capital across countries is to look at ratios of total stock of human capital to nominal GDP. The estimated ratios for 2006 are displayed in the right panel of Figure 3. On this basis, the ratios range between 6.9 in Norway and 11.7 in New Zealand, with an average value of around 9.5. Cross-country differences are much smaller when ignoring the four countries at both ends of the distribution. In fact, if we extend our estimation to include all available years for these countries, the estimated values of the minimum, maximum, mean and median are 6.9, 13.8, 10.0 and 10.0, respectively, with a standard deviation of 1.37.

33. For some countries, these values of the stock of human capital based on the methodology described in this paper can be compared with estimates derived from national studies that used the same methodology. Wherever possible we have tried to adapt our assumptions to generate new estimates in order to make the results as much comparable as possible. For instance, the comparisons are based on the same values for annual real income growth rate and discount rate as used by those for the corresponding national studies.

- In the case of Norway, the differences between the estimates based on the simplified methodology used here and those presented in Liu and Greaker (2009) range between 1% (in 2001 and 2006) and 21% in 2003, with an average difference (across the period) of 8%.
- In the case of Australia, the difference relative to the results in Wei (2004) is about -2% (in 2001).
- In the case of New Zealand, the difference compared with the results in Le et al, (2006) is higher, at about 30% for 2001.
- For Canada, our estimates of human capital per capita for total, male and female in 2006, the latest year available in our database, are 626,670, 799,230 and 452,530 (in Canadian dollar).

Future estimates of the stock of human capital could rely on country-specific values of these two parameters, possibly drawing on the OECD *Medium-term Baseline* which is a scenario postulated for projection analysis within the OECD.
respectively, while in Gu and Wong (2009) the corresponding estimates for 2007 are 652,600, 794,900 and 511,200 (in Canadian dollars).

These differences in numerical values between the estimates shown here and those produced by national research in the subject will be further investigated in the future.

**Figure 4. Gender distribution of human capital and gender ratio of human capital per capita, 2006**

![Gender distribution of human capital and gender ratio of human capital per capita, 2006](image)

Note: Estimates for Australia refer to 2001 and for Denmark to 2002.

34. The estimates of the stock of human capital based on the discounted life-time income approach can also be compared across groups within each country. While various types of decompositions are possible, Figure 4 shows the distribution of human capital between males and females (left panel), and the gender ratio of human capital per capita (right panel). In all countries, males account for around 60% of the total value of the human capital stock, with marginally lower shares in Romania and Poland, and marginally higher ones in Spain, Australia, Italy, Netherlands and Korea. Similarly, the ratio of human capital per capita between males and females ranges between less than 1.5 in Romania and Poland, and larger than 1.9 in Australia, Netherlands, Italy and Korea. Gender differences in lifetime labour income per capita reflect a combination of lower labour force participation, lower employment rate and lower pay for females than for males. As, in all countries, females do more housework than males, these gender differences in human capital per capita would be lower if estimates were extended to include non-market activities.

**Conclusions and next steps**

35. The preliminary estimates described in this report show the feasibility of applying the discounted lifetime income approach to measure the stock of human capital in several countries. They also highlight the feasibility of applying a simplified methodology based on the categorical data (e.g. by 5-year or 10-year age group) that are typically available within the OECD statistics system, rather than the more demanding detailed data that is required by the original methodology.

36. In the months to the end of 2010, the project will aim at improving the statistical information underlying the exercise (e.g. by improving the comparability of educational categories, by better understanding the nature of the earnings concept used in the OECD database, and by identifying possible alternatives) and extending the evidence produced to additional countries. Over this period, the project will also aim to derive estimates of the volume of the stock of human capital to illustrate changes over time; to conduct additional decompositions of the data at hand; and to analyse the sensitivity of results with respect to changes in a number of key parameters.
In the longer term, OECD work on human capital based on the accounting approach explored in this report could look at a number of more fundamental issues:

- Extending the scope of the current exercise to children (below age 15) and non-market activities, and projecting the value of human capital in the future, based on existing demographic projections.

- Constructing human capital accumulation accounts that will explain the changes in the stock of human capital over time in terms of investment, depreciation and revaluations.

- Exploring how estimates of adults’ competencies by attainment levels (based on PIAAC) could be integrated into this framework as ‘quality-adjustment’ of people’s attainment levels.

- Using these estimates of the human capital stock for the construction of an educational account integrating data on the various inputs entering into its production (monetary expenditures and non-market inputs) and of the outputs produced.

- Analysing the different ways in which human capital could improve the analysis of different economic aggregates and how the accounting approaches for these aggregates hang together. In particular, these concern the measurement of the output of the education sector, a set of extended household production accounts and the role of human capital in measuring quality-adjusted labour input for productivity analysis.
REFERENCES


Appendix

Figure A1: France (2006): Annual (AIN) and lifetime (LIN) income by educational level and age (Male) (National currency, current prices, in thousands)

Figure A2: France (2006): Annual (AIN) and lifetime (LIN) income by educational level and age (Female) (National currency, current prices, in thousands)

Figure A3: New Zealand (2006): Annual (AIN) and lifetime (LIN) income by educational level and age (Male) (National currency, current prices, in thousands)
Figure A4: New Zealand (2006): Annual (AIN) and lifetime (LIN) income by educational level and age (Female) (National currency, current prices, in thousands)

Figure A5: Norway (2006): Annual (AIN) and lifetime (LIN) income by educational level and age (Male) (National currency, current prices, in thousands)

Figure A6: Norway (2006): Annual (AIN) and lifetime (LIN) income by educational level and age (Female) (National currency, current prices, in thousands)
Figure A7: Poland (2006): Annual (AIN) and lifetime (LIN) income by educational level and age (Male) (National currency, current prices, in thousands)

Figure A8: Poland (2006): Annual (AIN) and lifetime (LIN) income by educational level and age (Female) (National currency, current prices, in thousands)

Figure A9: Spain (2006): Annual (AIN) and lifetime (LIN) income by educational level and age (Male) (National currency, current prices, in thousands)
Figure A10: Spain (2006): Annual (AIN) and lifetime (LIN) income by educational level and age (Female) (National currency, current prices, in thousands)

Figure A11: US (2006): Annual (AIN) and lifetime (LIN) income by educational level and age (Male) (National currency, current prices, in thousands)

Figure A12: US (2006): Annual (AIN) and lifetime (LIN) income by educational level and age (Female) (National currency, current prices, in thousands)