ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS

Working Paper 8
15 September 2009

ENGLISH ONLY

Joint UNECE/OECD/Eurostat Task Force for Measuring Sustainable Development
First meeting
Geneva, 23-24 September 2009
Session III of the Provisional Agenda

CONSTRUCTION OF HUMAN CAPITAL ACCOUNTS IN THE MEASUREMENT OF SUSTAINABLE DEVELOPMENT

Prepared by Barbara M. Fraumeni, Muskie School of Public Service, University of Southern Maine, the National Bureau of Economic Research, United States, and the China Center for Human Capital and Labor Market Research

I. INTRODUCTION

1. Discussions of sustainability typically focus on human actions that affect the well-being of current or future generations, notably though their impact on natural resources and the environment. Many authors cite a 1987 report of the World Commission on Environment and Development that defines sustainable development as development that meets "…the needs of the present without compromising the ability of future generations to meet their own needs." (World Commission on Environment and Development, 1987, p. 1) Until recently, little attention has been paid to human capital as itself a resource.[1] The focus has been instead on how humans determine the availability of other resources. But human capital plays a critical role in satisfying the needs and wants of humankind – through both market and nonmarket activities – and is thus an important part of any sustainable development story.

2. Human capital accounts can measure both the current and the potential future contribution of human capital to sustainability through economic growth.[2] If a population is growing, *ceteris paribus*, its economy must be growing to meet present and future needs. Needs can be met through market activities of humans, which both produce goods and services, and generate income that allows the acquisition of market goods and services. Those market activities ideally will result in adequate supplies of goods and services. Needs can also be met through nonmarket activities of humans such as the household production of cooking, cleaning, and care-giving. Some market income may be saved and invested, which allows for income and production to meet future needs. Some human capital activities, *e.g.,* education, and child bearing and rearing, are investments that increase future human capital. Without consideration of the human capital resources of a country, a sustainable development discussion is incomplete.

[1] A publication of the World Bank (2006, p. xiv) has highlighting the importance of human capital by concluding that worldwide intangible capital, which includes human capital and the quality of formal and informal institutions, is the largest form of wealth.

[2] Many important sustainable development issues will not be discussed in this paper, such as the impact of human and production activities on the environment and natural resources.

3. Human capital accounts are an obvious basis for recognizing human capital as a resource. Such accounts are an important component of any attempt to measure sustainable development.

4. Recently there has been a renewed interest in constructing human capital accounts. By 2005 a few researchers had adopted the approach of Jorgenson and Fraumeni (1987, 1992a, 1992b) in constructing human capital accounts. Alroth (1997) constructed Jorgenson-Fraumeni (J-F) accounts for Sweden, and Li, Gibson, and Oxley (2005) did the same for New Zealand. O'Mahony and Stevens (2004) used the J-F approach to estimate education output. More recently, several researchers undertook the development of J-F human capital accounts for other countries. A joint Fondazione Giovanni Agnelli/OECD human capital workshop in Turin, Italy, in November 2008, presented two papers using J-F methodology.[3] Gu and Ambrose's paper described human capital accounts for Canada, while Greaker and Liu presented a paper describing human capital accounts for Norway. At that same workshop, Fraumeni outlined a simplified approach to constructing J-F human capital accounts. Christian (2009) has estimated J-F human capital accounts for the United States for a recent period. Subsequently, researchers at the China Center for Human Capital and Labor Market Research undertook the construction of a set of J-F human capital accounts for China, and researchers at UK's Office for National Statistics began a similar project for the United Kingdom.[4]

5. Denmark, France, Korea, Mexico, and Romania are among the countries not listed above that have committed to the OECD human capital initiative, as well as two international organizations: Europstat and the International Labour Organization. Although it is not known if they would use J-F methodology in constructing human capital accounts, such accounts for those countries would be a net addition to already existing capital accounts.

II. J-F HUMAN CAPITAL ACCOUNT METHODOLOGY AND THE FRAUMENI SIMPLIFIED APPROACH

6. In 2006, I began to develop a simplified approach to constructing J-F human capital accounts (2006, 2008a, 2008b). I did so in recognition that national statistical agencies and other researchers might not be able to invest five years in constructing the accounts, which the original J-F approach required.

7. J-F created data sets by sex, individual years of age, and education (either highest level attained by individual level or enrollment by individual level) for the following categories: population, enrollment, employment, labor compensation, and hours of market work, in addition to survival rates by sex, and individual year of age. In some cases J-F began with incomplete or categorical data, but created all the required detailed data sets.

8. I call my simplified approach a categorical approach, a reference to the use of categorical data, *e.g.*, ages 25-34, instead of age 25, 26, 27,…35 in constructing the accounts. I conceptualized methodologies to deal with the possibility that all of the data sets listed above are categorical data.

---

[3] See the workshop web site at http://www.oecd.org/document/39/0,3343,en_2649_33715_41153767_1_1_1_1,00.html . (Retrieved September 5, 2009)

[4] On October 9-10, 2009 at a China Center for Human Capital and Labor Market Research international symposium in Beijing J-F human capital accounts will be presented for China. A paper abstract has been submitted by researchers from the UK's Office for National Statistics to present J-F human capital estimates at an International Association for Research in Income and Wealth conference in August of 2010 in St. Gallen, Switzerland.

9. All of the researchers writing papers in 2008 or 2009 use a partial categorical approach. In other words, for some data they had information by single year of age, single level of enrollment, and/or single level of educational attainment. Other data were available only by category. The researchers did not attempt to develop detailed level data; instead they assumed that category characteristics applied to all or some components of the category.

10. Conceptually, the J-F approach is not difficult. However, as Christian stated:

> "Human capital accounting is simultaneously one of the easiest and most difficult exercises in empirical economics. It is easy in the sense that the statistical techniques necessary are relatively simple. On the other hand, getting the data right can be massive challenge." Christian (2009, p. 1)

11. Getting the data "right" involves at least three challenges:

(a) Finding data;

(b) "adjusting" the data when they are not exactly what a researcher needs;

(c) Making reasonable assumptions regarding characteristics of detailed level data when only categorical data are available.

12. The most difficult problems are with the years during which individuals may be in school and, depending upon their age, also working. To complicate matters further, the number of years in an age category, *e.g.*, ages 6-15 and ages 16-17, often will not match the number of years that an individual is in an education level, *e.g.*, elementary school which takes six years to complete, and typically starts at age 6 in the United States, and high school which takes four years to complete and typically starts at age 14. But individuals may not complete an education level. And many individuals in the United States take longer than four years to complete a four-year college or university degree. Finally, not all individuals initially enroll in an education level at the same age.

13. It is probably a practical necessity to assume that individuals do not skip grades, fail grades, or return to school sometime after dropping out. Without making such assumptions, keeping track of where students are in the educational system becomes very complex.

14. Birth information may be particularly valuable to start progressing individuals through age categories when only categorical information is available, and by starting-age assumptions through school.

15. Ideally, categories should be at a sufficiently detailed level to at least help facilitate construction of the accounts. For example, it is unlikely that all individuals aged 6 to 24 have the same probability of enrollment. It might, however, be reasonable for researchers in the United States to assume that those aged 6-24,14 -15, 16-17, 18-21, and 22-24 have the same probability of enrollment by those more detailed age categories. Similarly, it is unlikely that all individuals aged 25-74 have the same probability of living another year. However, it might be reasonable to use more detailed age categories to assume that individuals aged 25-34, 35-44, 45-54, 55-64, and 64-74 have the same probability of surviving.

16. J-F requires the following data, by sex, individual years of age, and education (either highest level attained by individual level or enrollment by individual level): population, enrollment, employment, labor compensation, hours of market work, and survival rates.[5] All but survival rates are contemporary information.

17. Contemporary information and data sets are used to assess future probabilities that persons will go to school, perform market work (for how many hours), and live. Relative wage rates come from contemporary relationships. Future wage rates are assumed to increase at a specified rate.

18. Unlike recent researchers except for Christian, J-F estimated both market and nonmarket lifetime income.

19. The equation notation in the following equations which calculate the value of per capita lifetime income (human capital) is:

(a) Mi: lifetime market income ;

(b) Nmi: lifetime nonmarket income;

(c) Ymi: yearly (current) market income;

(d) Ynmi: yearly (current) nonmarket income;

(e) G: real rate of growth in labor income;

(f) R: discount rate;

(g) Sr: survival rate to one year older;

(h) s: sex;

(i) a: age, by single year of age, *e.g.*, age 0, 1, 2, ...74, 75+;

(j) e: highest level of education attained, by individual level of education from grade 1, 2, ..., through at least one year of graduate school;

(k) enr: grade level enrolled, by individual level of education, grade 1, 2, through at least one year of graduate school; and

(l) older: age + 1, *e.g.*, being one year older.

20. J-F allowed for five stages of life:

(a) Stage 1: No school or work, ages 0-4;

(b) Stage 2: School, but no work, ages 5-15;

[5] Hours of market work is needed only if nonmarket human capital accounts are constructed.

(c) Stage 3: School and work, ages 16-34;

(d) Stage 4: Work only, ages 35-74;

(e) Stage 5: Retirement, zero income, ages 75 or older.

21. The stages were determined by data availability, *e.g.*, even though some individuals go to school after age 34, until recently no data were available for the inclusion of those individuals.

22. All of the equations are estimated backwards recursively, *e.g.*, first lifetime income is estimated for those 75, then 74, 73,…0 (infants).

23. Equations for ages 35-74 (work only) are:

mi(s,a,E) = ymi(s,a,e) + sr(s,older) * mi(s,older,e) * (1+g)/(1+r);
nmi(s,a,e) = ynmi(s,a,e) + sr(s,older) * nmi(s,older,e) * (1+g)/(1+r).

24. Yearly nonmarket income is estimated by assuming 10 hours per day is spent either sleeping and in maintenance; that students spend 1600 hours per year in formal education activities; and that 14 hours per day is available to go to school, perform market work, and engage in nonmarket activities, including nonmarket work and leisure. Opportunity cost methodology is used to value nonmarket activities, and income is on an after-average marginal labor income tax basis.

25. Equations for ages 5-34 (work and school) are:

mi(s,a,e) = ymi(s,a,e) + sr (s,older) * [senr(s,a,enr) * mi(s,older,e+1)
+ (1 - senr(s,a,enr)) * mi(s,older,e)] * (1+g)/(1+r)
nmi(s,a,e) = ynmi(s,a,e) + sr (s,older) * [senr(s,a,enr) * nmi(s,older,e+1)
+ (1 - senr(s,a,enr)) * nmi(s,older,e)] * (1+g)/(1+r)

26. Equations for ages 0-4 (no school or work) are:

mi(s,a,E) = sr(s,older) * mi(s,older,e) * (1+g)/(1+r)
nmi(s,a,e) = sr(s,older) * nmi(s,older,e) * (1+g)/(1+r)

27. The volume of lifetime income (human capital) is created by a Divisia index, where the weights are the share of the value of lifetime income for the subcomponent in the value of total lifetime income, and the growth rates are population growth rates. For example:

growth rate of lifetime income volume(year,s,a,e) =
.5 * [share(year,s,a,e) + share(year-1,s,a,e)]
{ln[population(year,s,a,e)] – ln[population(year-1,s,a,e)]}

28. The lifetime income value equations are modified in the simplified categorical approach.

29. Categories for the simplified approach for the United States include:

Ages: 0-5, 6-15, 16-17, 18-24, 25-34, 35-44, 45-54, 55-64 and 65-74.

Education: 0(1)-8, 1-3 years of high school, 4 years of high school,
1-3 years of college, 4 years of college, and graduate school.
where those categories are for highest education level attained and enrollment level, and 0 refers either to less than 1-8 years of school completed, or not being enrolled in school.

30. In the revised equations, categorical data are used for age, highest education level attained, and enrollment data. In Fraumeni (2007, 2009) capital letters indicate the use of categorical data.[6]

III. EXAMPLES OF CATEGORICAL APPROACHES

31. The following examples give an idea of how researchers handle categorical data when the J-F methodology was premised on either having or producing data at the detailed level.

32. Gu and Wong (2009), who include ages 15 to 74 in their human capital measures, have population by category of highest educational level attained, but no information on enrollment level. They do have enrollment by sex and single year of age.

33. Their solution to that problem is to make assumptions about how old students are in each grade.

34. Their education categories are the following:

Category 1: 0-8 years of schooling;
Category 2: Some or completed high school (three years to complete);
Category 3: Some or completed post-secondary education (assumed to be a two- year program);
Category 4: Bachelors degree (four-year program)
Category 5: Masters degree or above (assumed to be a two-year program).

35. They assume:[7]

(a) Enrolled individuals aged 15 or older whose education attainment is category 1 have an equal probability of being enrolled in any particular year of education level category 2;

(b) Enrolled individuals aged 15 whose education attainment is category 2 are in the first year of education level category 3;

(c) Enrolled individuals aged 16 or older whose education attainment is category 2 have an equal probability of being in any particular year of education level category 3;

(d) Enrolled individuals whose education attainment category is 3 who:
Are less than or equal to 18 are in the first year of education level 4,
Are 19 are in the second year of education level 4,

[6] The simplified approach equations are not listed in this paper as they are identical to those for the full J-F approach, except that "a" becomes "A", "e" becomes "E", "enr" becomes "ENR", and "older" becomes "OLDER".
[7] Attachment to a February 11, 2009 email from Wulong Gu.

Are 20 are in the third year of education level 4,
Are 21 are in the fourth year of education level 4, or
Are 22 or older have an equal probability of being enrolled in any particular year of education level category 4.

(e) Enrolled individuals whose education attainment category is 4 who:
Are less than or equal to 22 are in the first year of education level 5,
Are 23 are in the second year of education level 5, or
Are 24 or older have an equal probability of being enrolled in any particular year of education level category 5.

36. In thinking about how to develop J-F human capital accounts with only categorical data, I realized that the Gu and Wong approach makes sense. However, from the detailed enrollment data I have from 1961 to 1986, I know that the initial age at which students enter grades 1-12 (through high school) varies to a greater extent than they assumed. For example about 95 percent of students entering an elementary or secondary grade level are of one age or another.

37. The specific ages in that example depend upon when age is measured during the year in question. About 50 percent in a grade are of the younger age (*e.g.* age 6 sometime in the year when entering grade 1) and about 45 percent in the same grade are of the older age (*e.g.* age 7 sometime in the year when entering grade 1). That knowledge of the underlying enrolled age distributions suggests modifying the Gu and Wong methodology to allow for two ages to be enrolled in a grade level.

38. Making a determination of the enrollment year of a multi-year category is particularly important because it determines how many years must pass before an individual can realize the lifetime income of a higher educational category.

39. With age and education information available only in categorical form, the Gu and Wong approach cannot be used directly without deriving population by single year of age from categorical data.

40. Nonetheless, population by single year of age can be imputed using the following procedure. The notation in the equations used to impute the individual age data is as follows:

B(s) is the number of persons born (of age 0)
Pop(s,1,1) is categorical population for age category 1 (ages 0-5) and education category 1 (grade 8 or less completed)
Population(s,a,1) is population by single year of age for education category 1 (grade 8 or less completed), and
Sr(s,1) is the average one-year rate of survival of individuals in age category 1 (ages 0-5)

41. Estimate population by single year of age and categorical level of education by using the following relationship:

$B(s) \rightarrow$ age 0 = population(s,0,1)
$Sr(s,1)*B(s) \rightarrow$ age 1 = population(s,1,1)
$Sr(s,1)^2*B(s) \rightarrow$ age 2 = population(s,2,1)
$Sr(s,1)^3*B(s) \rightarrow$ age 3 = population(s,3,1)

$Sr(s,1)^4 * B(s) \rightarrow$ age 4 = population(s,4,1), and
$Sr(s,1)^5 * B(s) \rightarrow$ age 5 = population(s,5,1).

42. It must be true after controlling that:

$\sum$ (n=0 to 5) population(s,n,1) = pop(s,1,1) or
$pop(s,1,1) = B(s) * \{1 + sr(s,1) + sr(s,1)^2 + sr(s,1)^3 + sr(s,1)^4 + sr(s,1)^5\}$
$= B(s) * \{\sum n = 0 \text{ to } \sum sr(s,1)^n\}$.

43. Solving for the number of persons born:

$B(s) = pop(s,1,1) / \{\sum (n = 0 \text{ to } 5)\ sr(s,1)^n\}$

44. After the equation is solved for B(s) the sum of the population(s,n,1) numbers are controlled to pop(s,1,1). Controlling is necessary as the survival rates come from a source that differs from the categorical population data.

45. There are two issues with that proposed methodology. First, survival rates are taken from the current year, as opposed to the year in which individuals survived to the next year. For example, if the number of individuals born in 1990 is being estimated, those from that birth cohort who survive are age 1 in 1991; 2 in 1992, …5 in 1995. The correct survival (to the next year) rates should be from 1990, 1991, …, 1994. Instead, to make that mathematically tractable, the survival rates are all taken from 1995. In the United States, survival rates change very slowly across years so it is not a problem. Even if survival rates were used from the correct years, controlling would still be necessary because survival rates do not come from the same source as the categorical population data.

46. The second issue is that the survival from age 0 to age 1 is significantly different from the survival rates from age 1 to 2, 2 to 3, 3 to 4, and 4 to 5. The difference arises typically from the health condition of infants at birth.

47. Drop-outs, which occur by age 16, if not before, in the United States, along with the fact that college and university students can take longer to finish school than the expected number of years, both present challenging issues.

48. Continuing to assume, as J-F did, that once a student drops out, he/she does not return to school at a later date is a convenient simplification. With that assumption, and the Gu and Wong approach, individuals can be tracked through the education system using the methodology described above for imputing detailed population.

49. Researchers at the China Center for Human Capital and Labor Market Research (CHLR) of the Central University of Finance and Economics faced another data challenge. In the absence of data on enrollment by detailed age and education levels, they employed data on initial enrollment by education level. The education categories are the following:

(a) Illiterate;

(b) Primary school (6 years);

(c) Junior middle school (3 years);

(d) Senior middle school (3 years); and

(e) College and above (3 years or more) or college (3 years) and university and above (4 years of more) starting in 2000.

50. They estimated the average probability that someone initially enrolled in an education level would advance to the next education level. Because of small sample survey properties, they used an average probability across all ages initially enrolling – rather than a probability specific to individual ages. The methodology then tracked how many students were at the different detailed grade levels to determine how many years it would take before students advanced to the next education level.

51. Average probabilities depended upon initial enrollments at the next higher education category level in subsequent years. An approach that might be considered to be more consistent with the J-F methodology would be to use contemporary information only. However, that approach was not adopted because advancement probabilities in China were increasing significantly over time. In the later years, advancement probability years were held constant at the latest available rate instead of predicting probabilities for 2008 forward.[8]

52. The CHLR researchers estimate labor income with Mincer equations.

53. Greaker and Liu (2008) have information by single years of age and by educational categories for enrollment and the labor force. They use the Norwegian Standard Classification of Education categories as the basis for their categories:

(a) No education and pre-school education;

(b) Primary education (seven years);

(c) Lower secondary education (three years);

(d) Upper secondary, final year (one+ years);

(e) Post-secondary non-tertiary education (one+ years);

(f) First stage of tertiary education, undergraduate (four years);

(g) First stage of tertiary education, graduate (two years);

(h) Second stage of tertiary education, postgraduate (one+ years);

(i) Unspecified.

54. They used data on years left to complete education to determine how many years pass before an individual advances to the next education level.

[8] I advised the researchers at CHLR on methodology.

55. Christian's approach (2009) is most similar to J-F. He faced a challenge in the switch of categories in the October Current Population Survey, the basic source: from one that asked about years in school, to one that asks about degrees or certificates completed.

56. He discusses the difficulty in separating aging and education effects: aging must occur at the same time as the increase in educational attainment.

57. A problem with the United States data – which may be present in data for other countries – is the "rose-colored glasses" effect. Data on population by highest education level attained are inconsistent with historical enrollment data. At least some individuals apparently report a higher level of educational attainment than they achieved. For example, given enrollment figures, educational attainment reported overall for the U.S. population is higher than could be possible.

IV. CONCLUSION

58. Although a variety of difficulties arise in constructing human capital accounts, recent progress indicates that many countries could develop such accounts. A common framework facilitates comparisons across countries when measuring human capital as a major factor in sustainable development. Given that J-F human capital accounts either exist or are being developed in eight countries, it would make sense to encourage adoption of J-F methodology as the basis for human capital measurement.

59. The examples given above of how different researchers have dealt with incomplete or aggregate level (categorical) data illustrate how researchers have successfully overcome data limitations. Although the data-availability situation of each country may differ, the variety of approaches that has already been used may well facilitate methodological development, and suggests promising techniques that could be adopted in pursuing a streamlined and simplified approach to constructing J-F human capital accounts.

V. BIBLIOGRAPHY

Ahlroth, Sofia, A. Bjorklund, A. Forslund (1997) "The Output of the Swedish Education Sector," Review of Income and Wealth Volume 43, Number 1, pp.89-104.

Christian, Michael S. (2009) "Human Capital Accounting in the United States: 1994 to 2006," paper presented at the Canadian Economic Association Annual Conference, May 29.

Fraumeni, Barbara M. (2006) "A Human Capital Approach to Measuring Government Education Output," presentation at the OECD/ONS/Government of Norway Workshop on the Measurement of Non-market Output and Health, London, England, UK, October 4.

_____ (2008a) "Human Capital and Investment in Education: A Streamlined Approach," presentation at the Fondazione Giovanni Agnelli/OECD Workshop on the Measurement of Human Capital, Turin, Italy, November 3.

_____ (2008b) "Human Capital: From Indicators and Indexes to Accounts," paper presented at the Fondazione Giovanni Agnelli/OECD Workshop on the Measurement of Human Capital, Turin, Italy, November 4.

_____ (2009) "Further Comments: Results by Country and Future Efforts" presentation at a China Center for Human Capital and Labor market Research seminar at the Central University of Finance and Economics, Beijing, China, June 19.

Greaker, Mads and Gang Liu (2008) "Measuring the Stock of Human Capital for Norway: A Lifetime Labour Income Approach," paper presented at the Fondazione Giovanni Agnelli/OECD Workshop on the Measurement of Human Capital, Turin, Italy, November 3.

Gu, Wuong (2008) "Reply to Barbara," attachment to a February 11 email.

Gu, Wulong and Ambrose Wong (2008) "Human Development and its Contribution to the Wealth Accounts in Canada," paper presented at the Fondazione Giovanni Agnelli/OECD Workshop on the Measurement of Human Capital, Turin, Italy, November 3.

____ (2009) "Human Development and its Contribution to the Wealth Accounts in Canada," paper presented at the Canadian Economic Association Annual Conference, May 29.

Jorgenson, Dale W. and Barbara M. Fraumeni (1989) "The Accumulation of Human and Non-Human Capital, 1948-1984," in R. Lipsey and H. Tice, The Measurement of Saving, Investment and Wealth, Chicago, University of Chicago Press, NBER, pp. 227-282.

_____ (1992a) "Investment in Education and U.S. Economic Growth," Scandinavian Journal of Economics, Vol. 94, supplement, pp. S51-70.

_____ (1992b), "The Output of the Education Sector," in Z. Griliches, T. Breshnahan, M. Manser, and E. Berndt (eds.), The Output of the Service Sector, Chicago, NBER, 1992, pp. 303-341.

Le, Trinh Van Thi, John Gibson, and Les Oxley (2005) "Measuring the Stock of Human Capital in New Zealand," Mathematics and Computers in Simulation, Volume 68, Issue 5-6, May, pp. 485-98.

O'Mahony, Mary and Philip A. Stevens (2004), "International Comparisons of Performance in the Provision of Public Services: Outcome Based Measures for Education," National Institute of Economic and Social Research, London, U.K.

The World Bank (2006) Where is the Wealth of Nations, Measuring Capital for the 21st Century, The International bank for Reconstruction and Development/The World Bank, Washington, DC.

Wei, Hui (2008) "Developments in the Estimation of the Value of Human Capital in Australia," paper presented at the Fondazione Giovanni Agnelli/OECD Workshop on the Measurement of Human Capital, Turin, Italy, November 3.

World Commission on Environment and Development (1987) "Resolutions Adopted by the General Assembly, 42/187. Report of the World Commission on Environment and Development," United Nations, Forty-second session, agenda item 82 e, December 11.

* * * * *