CHAPTER 4: Notes on the Construction of a Data Set for an O.E.C.D. Country

May, 2010

1. Overview

Our goal is to collect enough data on an OECD country (plus a few other countries that have good sources of data) so that we can construct a small applied general equilibrium model or macro model based on micro theory.

On the producer side, the model will have 5 commodities:

- Domestic output; an aggregate of consumption, government expenditures and investment;
- Exports;
- Imports;
- Labour input;
- Capital input.

Ideally, we would like to have land and natural resource inputs in the model as well but the old system of national accounts did not recognize the contributions of these inputs explicitly and thus historical data on these inputs are not available.

On the consumer side, we will not model the savings decision and so the consumer side model will have only 2 commodities:

- Consumption;
- Leisure demand (or the negative of labour supply).

We will also collect data on the various tax wedges that consumers and producers face.

We will collect data for the years 1960 to the latest year available for most OECD countries. Students can choose to collect data for a non OECD country but typically, it will prove to be difficult to find all of the necessary data (the student will have to use the data generated by the relevant national statistical agency). However, the Asian Productivity Organization has recently developed some very usable data sets for 6 non OECD Asian countries so these data can also be used. The data for recent joined OECD countries has typically not been pushed back to 1960. The list of OECD countries for which the data are available back to 1960 except where noted are as follows (I have excluded Canada since I will use Canada as an example in the files that I send you):

1. Australia;
2. Austria;
3. Belgium;
4. Denmark;
5. Finland;
6. France;
7. Germany (some complications here after the two Germanys unified in 1991);
8. Greece;
9. Iceland;

10. Ireland;
11. Italy;
12. Japan;
13. Korea (1970);
14. Mexico (1970);
15. Netherlands;
16. New Zealand;
17. Norway;
18. Portugal;
19. Spain;
20. Sweden;
21. Switzerland;
22. Turkey;
23. United Kingdom;
24. USA.

The OECD also has data for the following recently admitted countries to the OECD:

25. Chile (data go back to 1996);
26. Czech Republic (1990);
27. Hungary (1995);
28. Poland (1991);
29. Slovak Republic (1993);
30. Estonia (1995);
31. Israel (1995 but I believe that we can go back much further than this);
32. Slovenia (1995).

The OECD is also publishing data on some important non OECD countries (see OECD.STAT):

33. Brazil (data go back to 1990);
34. China (data back to 1970);
35. India (data back to 1997);
36. Indonesia (data back to 1990);
37. Russia (data back to 1995);
38. South Africa (data back to 1970).

The Asian Productivity Organization has usable data back to 1970 for China, Indonesia and India and the following three developing countries:

39. Republic of China (Taiwan);
40. Philippines;
41. Thailand.

The first 24 countries are lower risk in the sense that students in this class have successfully found enough data to the various country estimations that we will be doing later in the course.

The Asian Productivity Organization has just published the *APO Productivity Databook 2010*, Tokyo: The Asian Productivity Organization, which lists basic national accounts output data back to 1970 for the following countries: Bangladesh, Cambodia, ROC (Taiwan), Fiji, Hong Kong, India, Indonesia, Iran, Japan, Korea, Lao PDR, Malaysia, Mongolia, Nepal, Pakistan, Philippines, Singapore, Sri Lanka, Thailand, Vietnam, China and the US. They also list employment data. The main data that are missing have to do with commodity taxes and labour input but this information may be available from the UN national accounts or national statistical agency information.[1]

2. Basic National Accounts Data

Source 1: *National Accounts; Main Aggregates; 1960-1997;* Volume 1, 1999 Edition, Organisation for Economic Co-Operation and Development; Paris, 1999. Call Number: HC 79 I5 O751; *Location: Koerner Level 1 Stacks*. This is a thin volume. The most recent volume will be available at the Reference section of the Koerner Library, Level 2. The data for the years 1970 to 2009 are also available from the OECD online; see OECD.Stat. This is actually the most convenient source but the data does not always go back to 1960; i.e., it often starts at 1970 (or later for newer OECD countries).

Copy the data for your country that have the title: **Main Aggregates** or for the more recent publications, **Gross Domestic Product: Expenditure Approach**. We will be using the following series:

At current prices:
- Government final consumption expenditure;
- Private final consumption expenditure;[2]
- Increase in stocks;
- Gross fixed capital formation;
- Exports of goods and services;
- Imports of goods and services.

At constant price levels:
- Government final consumption expenditure;
- Private final consumption expenditure;
- Increase in stocks;
- Gross fixed capital formation;
- Exports of goods and services;
- Imports of goods and services.

In addition, we will use the following current dollar series that are listed under **Cost Components of the GDP** or in the more recent publications, listed as **Gross Domestic Product: Income Approach**:

---

[1] Students have successfully chosen Hong Kong, China and Taiwan as their country in the past.
[2] The System of National Accounts was revised in 1993 and so for the more recent data, Private final consumption expenditure has been replaced by two categories: (1) Household final consumption expenditure and (2) Final consumption expenditure of NPISH's (Non Profit Institutions Serving Households, such as trade unions, professional societies, political parties, religious organizations, sports, cultural or recreational clubs and charities.

- Indirect taxes;
- Subsidies;
- Consumption of fixed capital (depreciation);
- Compensation of employees;
- Operating surplus.

For each country, the above series are available for the years 1960 and 1969-1997 from the 1999 publication *National Accounts; Main Aggregates; 1960-1997;*Volume 1, 1999 Edition. Here is our first problem: how do we get data for the missing years? For the expenditure components of the GDP, (C+G+I+X−M), we can fill in for the missing years using some additional data series that are tabled in the same publication on pages 150-157. Tables 26-30 on these pages give us volume indexes for the 5 output components of GDP for each of the "old-time" OECD countries; these are essentially the constant dollar series divided by a constant. Tables 32-36 give us price indexes for these components of GDP. With the help of these series, we can fill in the data for the missing years, 1961-1968. Thus, please copy pages 150-157 in addition to the earlier pages that list the data for your specific country.

Source 2: *National Accounts; Main Aggregates; 1960-1989*, Volume 1, Organisation for Economic Co-Operation and Development; Paris, 1991. Call Number: HC 79 I5 O751*; Location: Koerner Level 1 Stacks*.

This volume will enable you to fill in the Cost components of the GDP (indirect taxes, subsidies, depreciation, compensation of employees and operating surplus) for the missing years, 1961-1968.

Here is where another problem can arise: namely, the data in this publication may not agree with the data in the Source 1 publication for the years 1960 and 1969. If the differences in your data are small for these two years in the two sources, just ignore the differences. However, if the differences are "large", try and adjust the data from the second source to blend in with the data from the first source. We assume that the data in a later publication is more reliable than the data from an earlier publication!

If you cannot find this volume, you can use other volumes of the same publication (which can be found at the Level 1 Stacks of the Koerner Library) to obtain information on the cost components of GDP for the missing years 1961-1968;.

Source 3: *National Accounts of OECD Countries; 1991-2002;*Volume 1, *Main Aggregates* or Volume IIa, *Detailed Tables*, Organisation for Economic Co-Operation and Development; Paris, 2004. Call Number: HC 79 I5 O751*; Location: Koerner Level 2 Reference*.

These two publications are the most recent publications for the OECD National Accounts. Information for the years 1991-2007 can be obtained from this publication. In order to obtain information for the most recent years, try to find the appropriate National Statistical Agency online and check whether the information for the last few years is available. Alternatively, almost everything from 1970 on is available online at OECD.Stat.

3. Labour and Population Statistics

The system of national accounts provides current and constant dollar estimates for the output components of a country (alternatively, current dollar estimates are provided along with price indexes or deflators for the current dollar estimates). However, until recently, constant dollar components for the inputs into production have not been routinely provided in the system of national accounts. In this section, we will focus on the problems involved in finding price indexes for labour input or finding estimates of constant dollar labour input.

From the tables on the Cost Components of GDP (see sources 1 and 2 above), we can obtain estimates of the annual compensation of employees; i.e., on the value of labour input for employees for each year. However, these current dollar values do *not* include the value of labour inputs from:

- unpaid family workers and
- self employed workers (or employers and persons working on their own account).

In recent years, for most OECD countries, the omission of unpaid family workers is not a big problem but for some countries, in the 1960's and 1970's, there was a considerable amount of unpaid family work (particularly in agriculture) and so we do have to make an adjustment to our data to allow for the input of these workers. With respect to the contribution of the self employed, during the past decade, the proportion of self employed workers has been increasing in many OECD countries so we cannot simply ignore the labour input of these workers. In the current system of national accounts, the value of the labour input of the self employed is part of *operating surplus*, which is the value of outputs less the value of employee labour input. Thus operating surplus includes capital depreciation, the return to capital (land rent, interest paid, dividends paid and an imputed return to equity capital employed in production) as well as an imputed payment for the labour services of the self employed and the unpaid family workers. We will make some rather arbitrary assumptions in order to obtain an imputation for the value of the labour input of the self employed and unpaid family workers such as assuming that these workers earn 70 % of the amounts that regular employees earn.

Source 4: *Labour Force Statistics: 1959-1970*, Paris: OECD, 1972. Call Number: HD 5764 A6 O58. Location: *Koerner Library, Level 1 Stacks.* (The same publication covering the years 1969-1989, 1980-2000 is in the same location). Thus data for all years 1960-2000 can be obtained at this location.

Please collect annual data for the following series:

- *Civilian employment* plus the *armed forces* which is *total employment*; total employment can also be calculated as the *total labour force* less the *unemployed*.
- *Wage earners and salaried employees* (or employees); these are the workers whose earnings are collected in Compensation of Employees in the cost components of the National Accounts;
- *Employers and persons working on their own account* (the self employed); these workers' compensation shows up in Operating Surplus;
- *Unpaid family workers*; these workers' compensation shows up in Operating Surplus.
- Population 15 to 64 years.

From the above sources, please collect data on the total population for your country as well as the number of persons from 15 to 64 years of age. Data for recent years is also available from the next source listed below.

Source 5: *Labour Force Statistics: 1982-2002*, Paris: OECD, 2002. Call Number: HD 5764 A6 O58. Location: *Koerner Library, Level 1 Stacks.*

There are several problems with taking the quantity of the labour input of employees to be proportional to the number of wage earners and salaried employees:

- these estimates make no allowance for changes in average hours worked by each employee in each year (over time, hours of work have tended to decline and more holidays and increased vacations have been offered to workers);
- these estimates make no allowance for the changing mix of full time and part time workers and
- each hour of work offered by workers of varying skills is regarded as being equal in its contribution to production.

In order to overcome the last problem listed above, the last version of the international system of national accounts (*System of National Accounts 1993*, Eurostat, IMF, OECD, UN and World Bank, Luxembourg, Washington, D.C., Paris, New York, and Washington, D.C.) recommended that countries construct a proper index number of wages (an employment cost index) but it is only in recent years that a few countries have actually implemented this suggestion.

Our suggested approach to solving the problem of obtaining an estimate of real labour input into the economy has been to use the number of workers as the estimator. However, another approach is to divide our estimate of the value of labour input by a price index, which leads to an implicit estimator for the quantity of labour input. We now pursue this second approach.

Source 6: *Main Economic Indicators: Historical Statistics 1960-1979*, Paris: OECD. Call Number: HC10 O68 H58. *Location: Koerner Library, Level 1 Stacks.* (The same publication covering the years 1962-1991 is in the same location). *Main Economic Indicators: Historical Statistics 1969-1988*, Paris: OECD. Call Number: HC10 O68 H58. *Location: Main Library, Level 1 Stacks.*

These publications contain annual indexes of either:

- weekly wage rates (all activities or just manufacturing) (New Zealand);
- hourly rates in manufacturing (France, Germany);
- hourly rates: industry (France);
- hourly earnings in manufacturing (Germany, Canada);
- unit labour costs in mining and manufacturing (Germany);
- unit labour costs in manufacturing (Canada).

It can be seen that different countries have different wage indexes. Try to pick the most comprehensive (annual) one that is available. Unit labour costs are generally preferable to hourly

or weekly wage rates because these series include fringe benefits and try to adjust for changes in the number of days worked each month or year. It may be necessary to link your best series for the years 1962-1991 with other series that are available for the remaining years. Again, there are problems associated with the use of these wage series as deflators for the compensation of employees:

- The wage series usually are not comprehensive; i.e., they cover only a portion of the economy (usually just manufacturing) and typically do not cover service industry wage rates.
- Usually, no adjustments are made (for the wage rate series) for changes in employee benefits (mainly pensions and medical coverage) or for changes in holidays or for changes in days paid for but not worked.

To complete your chosen wage rate series, use the following source:

Source 7: *Main Economic Indicators: (Monthly)*, Paris: OECD. Call Number: HC10 O68. *Location: Koerner Library, Level 1 Stacks.* (The same publication covering the earlier years is in the same location). Each of these monthly publications has the data only for 4 years so it will be necessary to use a number of these publications in order to complete your wage rate series for the years 1991-2008 (move forward only 3 years at a time so that you have an overlap year as you move forward). Note that the base year for these indexes (i.e., the year for which the index is set equal to 100) will vary from publication to publication. When the base year changes going from one publication to another, it will be necessary to collect data so that your series overlap each other for at least one year. Data for recent years are available from the following source:

Source 8: *Main Economic Indicators: (Monthly)*, Paris: OECD, March 2009. Call Number: HC10 O68. *Location: Koerner Library, Level 2 Reference*.

It should be noted that sources 6 to 8 can also be used to collect interest rates at the same time; see below.

An alternative source for an index of weekly wage rates (see the series: wages: hourly earnings) is:

Source 9: *International Financial Statistics*, Washington D.C.: The International Monetary Fund. Call Number: HG1 I55. *Location: Koerner Library, Level 1 Stacks.*

Source 10: *International Financial Statistics: Yearbook 2004*: Washington D.C.: The International Monetary Fund. Call Number: HG1 I552. *Location: Koerner Library, Level 2 Reference.*

4. Capital and Interest Rate Series

We will construct an annual series for capital input into the economy by using the data on investment or gross fixed capital formation collected in section 2 above (plus assumptions about the length of life of investment goods or assumptions about depreciation rates). However, we will also require a series on the price of financial capital or on the nominal interest rate that producers face.

It is somewhat difficult to find long time series on nominal interest rates that producers face so we will settle for a time series for government bond interest rates. Our source for this series is Sources 6 to 8 listed above. These publications contain a variety of interest rate series for various countries, including the following series:

- Yield of long term government bonds (New Zealand);
- Official discount rate (France, Germany);
- Call money rate (France, Germany);
- Bond yields: private corporations (France);
- Bond yields: issues guaranteed by the government (France);
- Public and semi public sector bonds (France);
- Treasury bill rate (these are very short term bonds of less than one year) (Germany);
- Yield of government bonds (Germany);
- 7-15 year public sector bonds (Germany);
- Federal government bonds (Canada);
- U.S. government bonds; composite over 10 years;
- Central government bonds (Japan);
- 10 year Commonwealth government bonds (Australia);
- Public sector bonds (Austria);
- Taxable public bonds; 3-6 years (Finland);
- Treasury bonds (Italy);
- Government bonds (more than two years) (Spain);
- Confederation bonds (Switzerland);
- 20 year government bonds (United Kingdom).

As can be seen from the above, there are a bewildering array of possible bond rates to choose from in some cases. Ideally, we would like to have a one year private sector bond rate but for most countries, such an ideal rate is not available going back to 1960. I would not use the official discount rate or Treasury bill rates; these series usually have quite different movements compared to other interest rates and hence these rates are not usually representative. From the remaining series, you will have to make your best judgement as to which series to use. It may also be necessary for you to link your earlier interest rate series taken from the *Main Economic Indicators: Historical Statistics* with a different series taken from the monthly publications. To see if two different series can be linked, check their behavior for the years that they overlap. If the two series are quite similar during the overlap years, then the two series can be linked into one composite series.

An alternative source for an interest rate series is Sources 9 and 10 listed above; see the "Government Bond Rate" series for your country. For Canada, for the early years, the listed series are:

- Bank rate (end of period);
- Treasury bill rate (these are short term government bonds);
- Government bond yield.

For the more recent years, for Canada, the listed series are:

- Bank rate (end of period);
- Money market rate;
- Corporate paper rate;
- Treasury bill rate;
- Deposit rate;
- Lending rate;
- Government bond yield, medium term;
- Government bond yield, long term.

My two preferred choices would be:

- The treasury bill rate or
- The government bond yield (early years) linked to the government bond yield, medium term.

5. Taxes and Tax Rates

The national accounts data on outputs that was collected in section 2 above includes all commodity taxes paid by the final demanders of the products and services. In particular, the price series include indirect commodity taxes that are paid for by purchasers. This is the correct treatment of commodity taxes from the viewpoint of a consumer, who faces the after tax prices on goods and services. However, from the viewpoint of producers, a price that includes commodity taxes that fall on outputs is *not* the price that producers face. The price of outputs that producers face should *not* include any taxes that are collected on the sales of outputs. Hence, when we are estimating production functions or profit functions (the producer side of our general equilibrium model), these indirect taxes that fall on outputs should be removed from output prices. In order to accomplish this removal, we require information on the amount of indirect taxes collected on outputs. A subsidy on the production of an output works in the opposite way to the way an output commodity tax works so what we require is information on indirect taxes on outputs less producer subsidies that fall on outputs. Now in section 2 above, we collected information on indirect taxes and subsidies so it would seem that the required information is readily at hand. However, there is a problem; namely, the series on indirect taxes includes not only the commodity taxes on outputs but it includes also certain commodity taxes that fall on *inputs* used by the production sector of the economy. Two examples of commodity taxes that fall on inputs are:

- property taxes on land and structures and
- gasoline taxes that fall within the integrated production sector of the economy; i.e., the energy taxes that trucks, railroads and aircraft pay while moving goods from industry to industry are taxes on the intermediate inputs used by the receiving sectors.

We will not model the second situation but it is necessary to remove property taxes from indirect taxes that fall on outputs. These property taxes can be better modeled as taxes that fall on the use of capital.

Another tax complication occurs with the treatment of labour. The wage rate that we have constructed in section 3 and the compensation of employees constructed in section 2 include (in theory) all of the income taxes that fall on labour income plus all of the payroll taxes paid by firms. These tax wedges should be removed in order to calculate the after tax wage rate faced by the consumer since it is the after tax wage rate that is relevant in modeling the labour supply decision. Thus while the price and quantity series that we have constructed for labour are the correct ones from the viewpoint of the firm, they are *not* the correct ones from the viewpoint of the consumer's utility maximization problem that generates commodity demand functions and labour supply functions.

The bottom line is that we need to allocate all of the taxes in the economy into 5 categories:

- taxes that fall on outputs;
- taxes that fall on exports;
- taxes that fall on imports;
- taxes that fall on labour and
- taxes that fall on capital.

There are two main sources that can help us with this tax allocation problem:

Source 11: *Revenue Statistics: 1965-1974*, Paris: OECD. Call Number: HJ2279 O75. *Location: Koerner Library, Level 1 Stacks.* At the same location, you can find the volumes for 1965-1980 (which has the data for 1965, and the years 1970-1979), 1965-1986 (which has the data for the years 1977-1985), 1965-1990 (which has the data for the years 1985-1990) and for 1965-1995 (which has the data for the years 1990-1994). The data for the more recent years, 1995-1999, can be found in *Revenue Statistics: 1965-2000*, Paris: OECD, 2001. Call Number: HJ2279 O75. *Location: Koerner Library, Level 1 Stacks.* Note that we cannot obtain data for the years 1960-1964 using these sources. However, aggregated data for 1960 can be found on page 185 of *Revenue Statistics: 1965-1990*, Paris: OECD. Call Number: HJ2279 O75. *Location: Koerner Library, Level 1 Stacks.* But OECD.Stat usually has all of the tax data except for the years 1961-1964. We will deal with this lack of data problem for these missing years in our econometric modeling by interpolating tax rates between 1960 and 1965. The data for the most recent years should be found in the following location (or online at OECD.Stat):

Source 12: *Revenue Statistics: 1965-200?*, Paris: OECD, 200?. Call Number: HJ2279 O75. *Location: Koerner Library, Level 2 Reference*.

However, all of these data can readily be downloaded using OECD.Stat.

Source 13: *Government Finance Statistics Yearbook*, *1980*, Washington, D.C.: International Monetary Fund. Call Number: HJ101 G67 H45. *Location: Main Library, Level 2 Stacks* or *Koerner Level 1 Stacks*. This volume has data for the years 1971 and 1973-1978. At the same location, you can find the volumes for 1988 (which has the data for the years 1977-1985), 1996 (which has the data for the years 1985-1994).

Source 14: *Government Finance Statistics Yearbook, 2001*, Washington, D.C.: International Monetary Fund. Call Number: HJ101 G67 H45. *Location: Koerner Library, Level 1 Stacks.* This volume has data for the years 1990-1999. At the same location, you can find the volumes for the years 1997, 1998 and 1999. The *Yearbook* for the most recent year has data for the years prior to the year of publication.

Using sources11 and 12, please table the following series:

- 1000, taxes on income, profits and capital gains;
- 1100, taxes on individuals (we will assume that these taxes fall on labour; we will assume that 1000-1100 falls on capital);
- 2000, social security contributions (we assume that these taxes fall on labour);
- 3000, taxes on payroll and workforce (we assume that these taxes fall on labour);
- 4000, taxes on property (we assume that these taxes fall on capital);
- 5000, taxes on goods and services (we assume that these taxes less customs and import duties fall on outputs);
- 5123, customs and import duties (we assume that these taxes fall on imports).

Using sources 13 and 14, please table the following series:

From Table A: Revenue and Grants, Consolidated Central Government:
- 1, tax on income, profits and capital gains;
- 1.1, individual (we will assume that these taxes fall on labour; we will assume that 1minus 1.1 falls on capital);
- 2, social security contributions (we assume that these taxes fall on labour);
- 3, taxes on payroll and workforce (we assume that these taxes fall on labour);
- 4, taxes on property (we assume that these taxes fall on capital);
- 5, domestic taxes on goods and services (we assume that these taxes fall on outputs);
- 6, taxes on international trade (we assume that these taxes fall on imports).

Other levels of government (state, regional or provincial governments): Revenue and Grants:

- 1, tax on income, profits and capital gains;
- 1.1, individual (we will assume that these taxes fall on labour; we will assume that 1minus 1.1 falls on capital);
- 2, social security contributions (we assume that these taxes fall on labour);
- 3, taxes on payroll and workforce (we assume that these taxes fall on labour);
- 4, taxes on property (we assume that these taxes fall on capital);
- 5, domestic taxes on goods and services (we assume that these taxes fall on outputs);
- 7, other taxes (assume that these fall on goods).

From Table L, Local Government, Part A, Revenue and Grants:

- 3, taxes on payroll and workforce (we assume that these taxes fall on labour);
- 4, taxes on property (we assume that these taxes fall on capital);

- 5, domestic taxes on goods and services (we assume that these taxes fall on outputs).

I have regarded export taxes as being equivalent to import taxes in the above. Most countries tax imports but do not tax exports so usually export taxes will be small.

It can be seen that the OECD *Revenue Statistics* are much simpler to work with than the IMF *Government Finance Statistics Yearbook*.