Economic Appraisal 6: Potential Pitfalls in the Economic Appraisal Process

The Economic Appraisal Process - and Potential Pitfalls

A. Introduction

The fundamental question that CBA addresses is:

Does a particular public sector investment (or policy initiative) constitute the best use of society's scarce resources?

To assess whether an intervention would constitute the best use of a society's scarce resources, two things are required:

1. The right information is assembled;
2. That information is processed and analysed correctly.

In Chapter 28 of his classic textbook 'Public finance: a normative theory' (1981, Business Publications Inc, Plano, Texas) the economist Richard Tresch consider these two maters, and in doing so considers several potential pitfalls that can plague the practice of economic appraisal and cost benefit analysis. Tresch's exposition makes clear that any considerations brought into a purported CBA study that are not germane to addressing this fundamental question are at best irrelevant and can often lead to conclusions being highly misleading (or just plain wrong!). He writes (page 556 of Public Finance) that:

> 'Many cost-benefit studies ...[emphasize] other factors that are essentially irrelevant to the fundamental question that cost-benefit analysis attempts to answer. ... Once these peripheral issues gain prominence, cost-benefit analysis loses its ability to discriminate among alternatives.'

This Topic considers a number of potential pitfalls, all of which one finds are rather common in economic appraisal in practice. In some of these cases, it is possible to say unconditionally what is a right and what is a wrong way of proceeding. In other cases, the correct way of proceeding is conditional on characteristics of the state of the economy in which the intervention is being considered. This conditionality requires that we exercise careful judgement in the choice of how to proceed in carrying out the EA/CBA.

B. Dangers of Incorrect Project Specification

In earlier Topics in this course, the importance of properly, fully and correctly specifying the Project, Programme, or Policy intervention being considered (or the set of options being considered) was stressed. Without doing so, the quality of any economic appraisal will inevitably suffer.

Of equal importance is the need to give a precise statement of the baseline or counterfactual against which the Project (or Project options) is (are) being appraised.

An intuitive understanding of why this involves is obtained by thinking in terms of proposed macroeconomic policy interventions. The analyst here will typically have available a dynamic macroeconomic model of the economy in question. The analyst will input into this model a set of values for policy variables (such as those related to the level and structure of government spending and taxation) that are expected to pertain under a "business as usual" or "minimum change from where we are now" baseline. The model will then be solved to establish the time paths of a set of state variables of interest (such as inflation, unemployment, and GDP growth rates) under this so-called **Base Run** of the model.

Next the modeller will consider some policy change that has been mooted. He or she will change the values of policy variables or parameters accordingly, and will then re-solve the model with the new set of policy variable values. The resulting time paths of the state variables constitute the **Simulation Run** of the model. Finally, calculating one further time path - the difference between the Simulation Run and the Base Run of the macromodel - we obtain data about the net impact of the policy intervention, from which the desirability of the policy intervention can be appraised.

This is, in essence, what all economic appraisal (EA) does. Hence, in doing any EA, if one improperly specifies the base run or the simulation run the output obtained will not be a proper basis for appraisal of the intervention in question.

Let us flesh things out with an (hypothetical) example, this time taken from the world of CBA rather than macro-modelling. Suppose that we wish to appraise a proposal to construct a hydro-electric power (HEP) generation project, based around the construction of a large dam. Initial thinking by the appraiser and his or her team suggest that the costs of the HEP project include the capital costs of constructing the dam and its HEP plant; the operating and maintenance costs of the facility; and its dismantling and decommissioning costs. Similarly, on the benefits side, the appraisal team conjectures that the HEP will:

- produce electricity;
- avoid producing carbon emissions (by virtue of using HEP rather than carbon technology);
- conserve scarce fossil fuels and other cost inputs that fossil-based energy would entail.

We shall suppose that the project takes five years to construct, and thereafter produces power for 46 years. The graphic below describes initial estimates of the various annual cost and benefit flows (in units of £ '000). Note that we use the conventions that

1. all benefit and cost flows take place on the first day of any year, and
2. that the first year of the project is year 0

The project, therefore entails construction during years 0 to 4 and power generation during years 5 to 50. The graphic is a screen shot from an Excel workbook available at the following link, the first sheet of which lists estimates of the monetary values of relevant benefit and costs flows, and from which all further analyses we discuss below are obtained.

Excel file: Counterfactual_v1.xlsx

	A	B	K	L
1				
2		Value of HEP Power Produced Per Year	B1	1150
3		Value of Carbon Emissions Avoided per year (by using HEP rather than fossil energy)	B2	380
4		Value of fossil fuel and other operational costs per year if power produced by coal plant	B3	650
5				
6				
7				
8		Initial Capital Costs (annually for first 5 years, Years 0 to 4)	C1	2200
9		Operating and maintenance costs of HEP scheme, per year	C2	400
10		Decommissioning costs (Year 51)	C3	1000
11				
12				
13		Discount rate (real terms)	r	0.05
14				

Let us look at one (flawed) attempt to compute the NPV of this project. The graphic below shows the worksheet in question. Note that the project appears to be substantially net beneficial, with an NPV of £16.1 million and a BCR of 2.0.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1																	
2				Benefits					Costs					Undiscounted		Discounted	
3		Year		B1	B2	B3	B		C1	C2	C3	C		NB		DNB	
4																	
5		0					0		2200			2200		-2200		-2200	
6		1					0		2200			2200		-2200		-2095.24	
7		2					0		2200			2200		-2200		-1995.46	
8		3					0		2200			2200		-2200		-1900.44	
9		4					0		2200			2200		-2200		-1809.95	
10		5		1150	380	650	2180			400		400		1780		1394.677	
11		6		1150	380	650	2180			400		400		1780		1328.263	
12		7		1150	380	650	2180			400		400		1780		1265.013	
13		8		1150	380	650	2180			400		400		1780		1204.774	
14		9		1150	380	650	2180			400		400		1780		1147.404	
15		10		1150	380	650	2180			400		400		1780		1092.766	
16		11		1150	380	650	2180			400		400		1780		1040.729	
17		12		1150	380	650	2180			400		400		1780		991.1706	
18		13		1150	380	650	2180			400		400		1780		943.972	
19		14		1150	380	650	2180			400		400		1780		899.021	
20		15		1150	380	650	2180			400		400		1780		856.2104	
21		16		1150	380	650	2180			400		400		1780		815.4385	
22		17		1150	380	650	2180			400		400		1780		776.6081	
23		18		1150	380	650	2180			400		400		1780		739.6268	
24		19		1150	380	650	2180			400		400		1780		704.4064	
25		20		1150	380	650	2180			400		400		1780		670.8633	
26		21		1150	380	650	2180			400		400		1780		638.9174	
27		22		1150	380	650	2180			400		400		1780		608.4928	
28		23		1150	380	650	2180			400		400		1780		579.5169	
29		24		1150	380	650	2180			400		400		1780		551.9209	
30		25		1150	380	650	2180			400		400		1780		525.6389	
31		26		1150	380	650	2180			400		400		1780		500.6085	
32		27		1150	380	650	2180			400		400		1780		476.77	
33		28		1150	380	650	2180			400		400		1780		454.0667	
34		29		1150	380	650	2180			400		400		1780		432.4445	
35		30		1150	380	650	2180			400		400		1780		411.8519	
36		31		1150	380	650	2180			400		400		1780		392.2399	
37		32		1150	380	650	2180			400		400		1780		373.5618	
38		33		1150	380	650	2180			400		400		1780		355.7731	
39		34		1150	380	650	2180			400		400		1780		338.8315	
40		35		1150	380	650	2180			400		400		1780		322.6967	
41		36		1150	380	650	2180			400		400		1780		307.3302	
42		37		1150	380	650	2180			400		400		1780		292.6954	
43		38		1150	380	650	2180			400		400		1780		278.7575	
44		39		1150	380	650	2180			400		400		1780		265.4834	
45		40		1150	380	650	2180			400		400		1780		252.8413	
46		41		1150	380	650	2180			400		400		1780		240.8013	
47		42		1150	380	650	2180			400		400		1780		229.3345	
48		43		1150	380	650	2180			400		400		1780		218.4138	
49		44		1150	380	650	2180			400		400		1780		208.0132	
50		45		1150	380	650	2180			400		400		1780		198.1078	
51		46		1150	380	650	2180			400		400		1780		188.6741	
52		47		1150	380	650	2180			400		400		1780		179.6896	
53		48		1150	380	650	2180			400		400		1780		171.133	
54		49		1150	380	650	2180			400		400		1780		162.9838	
55		50		1150	380	650	2180			400		400		1780		155.2226	
56		51									1000	1000		-1000		-83.0512	
57																	
58						PV(B) =	32,067.75					15,968.13					
59														NPV =		16100	
60														BCR =		2.008	

But a moment's consideration shows that this cannot be correct. Either

1. the HEP project is ADDITIONAL TO existing electricity capacity (in which case we should include the value of the additional power but we should not include the value of coal savings nor the value of carbon emissions avoided, as no less coal is being used in the economy), or
2. the HEP project is A SUBSTITUTE FOR existing electricity coal-fired capacity (in which case we should include the value of coal savings and the value of carbon emissions avoided, as less coal is being used in the economy, but we should not include the value of additional power, as there is none).

Looked at another way, the flawed analysis follows from a failure to correctly specify the counterfactual. There are two possible counterfactuals:

Counterfactual 1 (CF1)	The HEP power station replaces existing coal fired stations of same electricity output value
Counterfactual 2 (CF2)	All power generated by HEP station is additional to existing electricity power capacity(and is used to meet supply shortfall relative to demand)

By including all of B1, B2 and B3 as benefits, the flawed NPV analysis in effect assumes that the HEP project is **both completely additional to** existing power capacity **and (at the same time) in full a substitute for** part of that existing capacity. Clearly, this cannot be the case.

Mixed cases are possible, where a new project is used **partly** to replace existing capacity and **partly** to add to capacity. But in such a mixed case, the NPV of the mixed project would need to properly apportion the various benefits according to the degree of mixing involved. Thus, not the full totals of the set of three benefits could be attributed to the mixed project. (See the Exercise below).

You should download and then extend the workbook Counterfactual_v1.xlsx and then calculate the NPV for the two cases where the counterfactual is CF1 and where it is CF2. It would also be useful to supplement these two NPV calculations with that for a third special case, CF3:

Counterfactual 3: In CF3, one half of the HEP electrical capacity is used to replace existing coal-generated electricity, and the other half provides additional electrical capacity.

When you have done so, please briefly report your results and discuss the implications of your findings by a post to the Discussion Forum: NPV calculations for the HEP scheme under counterfactuals CF1, CF2 and CF3 .

C. Dealing with Inflation

Practitioners are often in doubt about how one should treat, if at all, general changes in the price level - that is, inflation - in economic appraisal. In the notes below we deal with the special case of Cost Benefit Analysis CBA), but the arguments apply to all forms of economic appraisal.

This is one area where the economic theory is clear: the practitioner must use one or other of the following methods (but must not mix the two):

METHOD 1	Express all future values that are used as inputs into the discounted cash flow (DCF) analysis in **REAL (CONSTANT PRICE)** terms and, when discounting, use an appropriate **REAL** rate of return. This amounts to assuming that the rate of inflation over the lifetime of the project is zero; the overall general price level remains unchanged over the project lifetime. The NPV of the 'project' is then the sum of the (constant price) discounted cash flows.
METHOD 2	Express all future values that are used as inputs into the DCF analysis in **NOMINAL (CURRENT PRICE)** terms and, when discounting, use an appropriate **REAL** rate of return. (Note that we discount again by a real rather than a nominal discount rate: this is not a typo.) But a further step is now required. As the discounted cash flows will now be in nominal (rather than constant) price terms, these must now be converted into constant price discounted cash flows by deflating each discounted flow by a factor that is equal to the proportionate increase in the average price level between the initial year and the year in question). The NPV of the 'project' is then the sum of these constant price discounted cash flows

These two methods give identical answers in terms of the NPV of the project. In that sense, as both give the same correct answer, it does not matter which one uses.

To demonstrate that the two rules are indeed exactly equivalent, please study the Excel file DCF.xls linked here. The file has been set up so that you can change any of the input data, and see that the equivalence will remain valid.

But clearly it is far simpler to use the former method, Method 1, which works in real (constant price terms), and so assumes that the rate of inflation will be zero over the project lifetime. It is simpler for three reasons:

1. The number of calculations is fewer
2. It is more transparent
3. It avoids the need to predict inflation rates over the lifetime of the project, a task that will almost certainly be nearly impossible to do with reasonable accuracy in practice.

As a matter of routine, therefore, you should use Method 1.

The real and nominal discount rate

All that remains is to explain the relationship between a real and a nominal rate of interest (or discount rate), and how one can be converted into the other.

First, we outline some general principles:

- *The nominal rate of interest: this is the rate of interest expressed in terms of some currency (also called the "current price" rate of interest)*
- *The real rate of interest: this is the rate of interest expressed in terms of purchasing power (also called the "constant price" rate of interest)*
- *The real rate of interest is equal to the nominal rate of interest corrected for the effects of inflation (i.e. changes in the purchasing power of units of the money currency in question).*

How do we obtain the real rate of interest (let us call it here the real rate of return)? An approximate answer is given by

> real rate of return = nominal rate of return **minus** inflation rate over period in question

An exact answer involves the formulae below (where rates of return are expressed in proportionate rather than percentage terms):

$$1 + \text{Real rate of return} = \frac{1 + \text{Nominal rate of return}}{1 + \text{Rate of inflation}}$$

or equivalently

$$\text{Real rate} = \frac{\text{Nominal rate of return - Rate of inflation}}{1 + \text{Rate of inflation}}$$

D. Pure Double Counting

Pure double counting occurs when an economic appraisal inadvertently counts the same item of benefit or cost twice in the process of carrying out the appraisal. This appears to have been a common mistake in the early years of carrying out CBA analyses, but is probably less common today. Nevertheless, it is a pitfall into which the appraiser can easily fall, and so care needs to be taken to avoid it.

A good example of double counting is given by Tresch (1981, page 561), which in turn is based on a case studied by McKean (1958, "Efficiency of Government Systems through Systems Analysis"). The example relates to the benefits of land that is irrigated by means of publically funded and projects. In 1958, the *Manual* of the Bureau of Reclamation within the U.S. Department of the Interior required that the benefits of irrigated land be counted as the sum of

1. The increase in the value of the land
2. The present value of the stream of net income obtained from farming the irrigated land.

Public irrigation projects are beneficial technological externalities that change the land's production function. The irrigation of previously non-irrigated agricultural land does deliver (gross) social benefits, provided that the irrigation does increase agricultural yields. But the question is should both items listed above be used as a measure of those benefits?

The answer is that they should not. Only one should be included. Including both is an instance of pure double-counting. A moment's reflection shows that 1 and 2 above are each a reflection of the same thing. An irrigation project that raises farm output will yield a higher net income flow in any period. Item 2 above - the present value of those flows - is capital value of set of increased net incomes. In the presence of a well-functioning price system, that increase in the wealth of the land will be reflected in a correspondingly higher market value, that is Item 1.

Tresch shows a second way of understanding why the official advice in this instance constituted a requirement to double count. A landowner can only benefit from one or the other of these two values, not both. The farmer could either continue to farm the land, thereby capturing the resulting stream of higher net incomes. Or the farmer could sell the land, capturing the increase in the land's value. It is not possible to capture both.

The example we have considered here is a case of 'pure' double counting where the same item of benefit (or cost) is being counted twice. Conceptually, it is equivalent to the mistake that would be made if one estimated GDP as the sum of the estimates of GDP obtained from the income, output and expenditure methods, rather than as just one of them (although that would be triple rather than double counting).

The term double counting is sometimes used in a more loose sense than the pure double counting considered on this page, and it is related to the issue of whether 'secondary' or 'indirect' costs or benefits should be included in an economic appraisal. As this question raises more fundamental issues we shall defer its consideration until the next page.

E. Dealing with secondary, indirect or induced effects

E1. What are meant by secondary, indirect or induced effects of a project?

Any project is likely to have a variety of indirect and/or induced impacts, in addition to the direct impacts of the project itself. Sometimes the word 'secondary' is used when referring to such indirect or induced effects.

These secondary impacts can take several forms, including the following:

1. **Changes in profits in industries directly involved in the project and in those which are indirectly affected by the project.**

   The project will almost certainly generate changes in profits for firms (and industries) directly and indirectly affected by the project. In terms of direct impacts, firms involved in the provision of the project itself will tend to experience rising profits. In terms of indirect (or secondary) impacts, profits will tend to rise in activities that are complementary to the project, and fall in those that have a substitute relationship to the project.

2. **Impacts on output and employment.**

   The direct impacts of the project will entail changes in resources used and/or goods and services produced. But because of the technological interconnectedness of productive sectors in an economy, those **direct impacts** will be associated with further **indirect changes** in output and employment in other sectors of the economy. Note, also, that if the project substitutes for other activities, or the consequences of its funding crowds out other activities, then indirect impacts can work in both directions.

3. **Induced impacts arising from changed household incomes.**

   A project will alter the magnitudes and distribution of incomes and wealth in a economy. For, example, those employed in a project may have higher incomes than they had before the project. When those higher incomes are spent, there will be further '**induced**' impacts on output and employment in other related sectors of the economy. But note again that a project can also lead - via substitution or crowding out effects - to falling incomes elsewhere in the economy.

4. **Consequences of changes in relative prices.**

   Any project that is of significant size (i.e. non-marginal) will change relative prices throughout the economy. These changes in relative price ratios will induce substitution effects in factor and/or product markets, with potential implications for economic welfare.

E2. Implications of secondary, indirect or induced effects for economic appraisal

These considerations beg an important question: to what extent should all, or some of, these indirect and/or induced impacts be included in the benefit or cost flows being used in an economic appraisal? This is a very big question, and one for which it is not easy to provide unequivocal answers. What is clear, though, is that the answer will have potentially huge implications not only for the difficulty of doing economic appraisal but also for the likely outcome of any particular appraisal. Tresch illustrates why the outcome can depend on what is and is not included in a project's costs and benefits, by considering the case of a hydroelectric project. The project which he has in mind is one which produces electricity (and so either makes this product available to some people or firms who previously had no access to it, or reduces the price of electricity to firms and/or households) and which will also provide additional services in the form of flood control, irrigation, and recreational facilities. All of these are true benefits of the dam, and must be properly estimated and evaluated as project benefits. But proponents of dams (and indeed almost any kind of project) often argue that there are additional impacts which should be included in an economic appraisal. These additional benefits are variously called indirect, induced or secondary benefits. Tresch writes that:

'Presumably the lowered price of electricity is the primary purpose of the project, but these other benefits (flood control, irrigation and recreational) are legitimate technological external economies created by the dam, and a careful analysis should try to evaluate them as part of the benefits. Proponents often go far beyond this, however, claiming as legitimate benefits profits arising in other industries as a consequence of the dam. ... With all these "secondary" profits added to the other benefits of the project listed above, the hydroelectric project is virtually assured of having a positive present value." (Tresch, 1981, page 557 558-559.]

A little later (page 558-559) he goes on to write:

"... suppose the damsite were formerly a wilderness area. Presumably the construction and continued operation of the dam will support all sorts of ancillary services. The people associated with the project have to be clothed, housed and fed. Indeed, a small town might spring up around the damsite, generating a continual flow of income in a region formerly devoid of economic activity. By the very nature of the Keynesian multiplier, these "secondary" income effects will be a multiple of the project's direct costs and/or benefits. Therefore, if they are included as project benefits, the project would necessarily have a positive present value. .... Something must be wrong here, for it appears to suggest that any project placed in an underdeveloped region of a country will be worthwhile."

Before we consider what might be wrong here (and why inclusion of increased profits is likely to be inappropriate), let us just note that Tresch's damsite example contains impacts of the first three categories listed in the previous section, and probably the fourth category too.

E3. Some basic principles that can inform the appraiser how to proceed

We can use some basic principles covered in earlier topics of this course to serve as criteria for helping us decide what kinds of indirect or secondary benefits might validly be included in an economic appraisal.

Principle 1: The worthiness of a project is to be judged by its impacts on social welfare

The first, and most important principle, is that the worthiness of a project is to be judged by its impacts on social welfare. To keep matters simple, let us assume for the moment that the project which is being appraised is a single project, and that it is being appraised relative to the counterfactual that the project does not take place. Ideally, therefore, one would like to know, **for each and every member of society**:

1. His or her utility (or some income-based proxy for that utility) in every period over the lifetime of the proposed project **under the assumption that the project does not take place** (i.e. under the counterfactual)
2. His or her utility (or some income-based proxy for that utility) in every period over the lifetime of the proposed project **under the assumption that the project does take place**. (i.e. under the counterfactual)

This exercise corresponds to that dealt with in section 2.5.2 (of Topic 5) where we discussed the interpretation of CBA as a welfare increase test. (Recall that this assumed that utilities can be aggregated via social welfare functions.) Table 11.7, which we reproduce below, deals with the special case of project with a 4 period lifetime and a society consisting of just 3 individuals. The notation used is such that $\Delta U_{B,2}$ denotes the change in utility during time period 2 that would be experienced by individual B on account of the project if it went ahead. It is the difference between the timestreams of utility under 2 and 1 above for individual B in period 2. Then for a generally agreed social welfare function with dated individual (cardinal) utilities as arguments, the analyst could compute

$$\Delta W = W(\Delta U_{A,0} , \ldots , \Delta U_{C,3})$$

and consider its sign; if positive the project should go ahead. We also noted in Topic 5 that alternatively, we could imagine that there existed an **intra-temporal** social welfare function which mapped individual utilities into a social aggregate, ΔU_t, in each period, and an **inter-temporal** social welfare function for aggregating over time. The analyst would then compute

$$\Delta W = W(\Delta U_0, \Delta U_1 , \Delta U_2 , \Delta U_3)$$

and the decision would be based on the sign here.

Table 11.7 Changes in utility (ΔU) consequent on an illustrative project

	Time period				
Individual	**0**	**1**	**2**	**3**	**Overall**
A	$\Delta U_{A,0}$	$\Delta U_{A,1}$	$\Delta U_{A,2}$	$\Delta U_{A,3}$	ΔU_{A}
B	$\Delta U_{B,0}$	$\Delta U_{B,1}$	$\Delta U_{B,2}$	$\Delta U_{B,3}$	ΔU_{B}
C	$\Delta U_{C,0}$	$\Delta U_{C,1}$	$\Delta U_{C,2}$	$\Delta U_{C,3}$	ΔU_{C}
Society	ΔU_{0}	ΔU_{1}	ΔU_{2}	ΔU_{3}	

As noted earlier, given that individual utility variations are not generally regarded as something that could be estimated *ex ante*, or observed *ex post*, the appraiser would in practice have to take an individual's utility to be a function of her total consumption, and to equate individual net benefit to the change in an individual's total consumption. If a project causes an individual to suffer a reduction in utility, that loss is expressed in monetary terms, and treated as a consumption loss for the individual. Similarly for gains. Adding across losses and gains for the individual gives her total consumption change, or net benefit, due to the project. We also noted in the previous topic that distributional considerations would imply that, in each period, contemporaneous total net benefit should be defined as the weighted sum of individual net benefits, with 'presumed' marginal utilities of consumption as weights, rather than as the simple sum. That is, using

$$NB_t = U^A_C \, NB_{A,t} + U^B_C \, NB_{B,t} + U^C_C \, NB_{C,t}$$

where U^i_C is the *i*th individual’s marginal utility of consumption, instead of

$$NB_t = NB_{A,t} + NB_{B,t} + NB_{C,t}$$

If this exercise were feasible, it would tell us how a project would affect very individual or household in the community in question, and so give us all the information we would need to decide whether or not the project were welfare enhancing. In a world of full information, or one where all required information could be obtained at zero cost, that kind of exercise is what should be carried out.

But of course, this exercise is not feasible in its entirety in practice. It would require identifying:

1. each individual, or groups of individuals, affected by the project;
2. by how much each individual's utility (or consumption possibilities) would be changed by the project were it to take place;
3. ascertaining the marginal utilities for the affected individuals or groups, so that an appropriate weighting can be decided upon.

Given the fact that impacts will be both direct and indirect (or secondary), and in a world of imperfect information or uncertainty, the best that is possible is to make estimates of each of items 1 to 3 above. Moreover, the amount of effort that is economically justified in searching for relevant information will be influenced by the magnitude of the project in question.

Two tools that may help in identifying secondary impacts of a project.

Input-Output (IO) Analysis

A key tool used by economists to identify and estimate the magnitude of these indirect changes is Input-Output analysis. Input-Output analysis conventionally sets out to estimate the impacts of changes in final demand on an economy. A change in final demand can have two kinds of further impact:

- indirect effects (arising from technological relationships between sectors of an economy)
- induced effects (arising from changes in household income that result from the change in final demand)

IO analysis distinguishes formally between indirect and induced impacts of changes in final demand by means of different multipliers that it examines:

- Type 1 multipliers include only the indirect impacts of the direct change in final demand.
- Type 2 multipliers include both indirect and induced impacts of the direct change in final demand.

This terminology is also useful as it suggests that IO Analysis can be a useful way of analysing the impacts of projects that can be legitimately thought of as directly leading to changes in final demand. But where a Project entails supply-side changes to an economy, the tool of IO analysis will, in most circumstances, be inappropriate.

A major problem with using IO analysis in economic appraisal work is that it is based on two key premises, both of which may not be valid for the project or policy being appraised. The first is that IO modelling assumes that the supply side of the economy is completely passive, in the sense that factor usage and output rise or fall to accommodate whatever level of demand there is in the system. IO modelling is entirely demand-driven. Clearly if there is the economy is in a state of generalised excess supply (underemployment of all productive factors in the economy - all factors, note, not just labour) then supply may be able to passively adjust to demand. But this is not likely to be true in general. The second is that production takes place through fixed coefficient Leontief technology, in which output is produced using fixed proportions of factor inputs. Hence a rise in final demand will bring about only positive changes in factor uses and in output. There is no scope for counteracting changes in output, employment and incomes elsewhere, as substitution effects and crowding out effects are excluded by assumption. Thus uncritically running a project's estimated change in expenditure through an IO model may well give grossly misleading estimates of its net overall impacts.

Computable General Equilibrium (CGE) Analysis

Where a Project directly alters the supply-side of the economy (such as through reforms to the labour market or the education or training sectors), relative prices of factor inputs, and of goods and services, are likely to change. If we now drop the assumption of fixed proportions Leontief technology, and allow for productive inputs to be substitutable for one another to some degree, then both demand side and supply side changes to an economy can be modelled using computable general equilibrium (CGE) analysis techniques that can allow for changes in relative prices (and so a variety of substitution effects) in the economy.

A CGE model could thus in principle be able to trace out all impacts - positive and negative, direct and indirect - of any policy intervention. But in practice, data limitations and computational feasibility limit the degree if sectoral disaggregation that can be accommodated by a CGE model, and so while its results might be indicative of broad sectoral changes, they certainly could not give us data at the individual level that a literal reading of Table 11.7 would imply. For readers interested in learning more about CGE analysis, please follow the CGE button link at the foot of this page. (The notes there are extracted from the Strathclyde University Open Learning MSc EMP degree class in Macroeconomics. They are quite extensive and are relatively technical in places. So this is very much **VOLUNTARY** reading for those particularly interested in this particular topic. Participants should note that these CGE pages can be omitted with no loss of continuity as far as this Economic Appraisal course is concerned.)

How does Principle 1 help us to understand whether secondary profit impacts should be included? The key point here is that the appraiser must not be selective. If changes in profits are to be included in the project evaluation, then **all** changes in profits in all areas of economic activity beneficially and adversely impacted by the project must be included, not just those that increase as a result of the project. This seems to suggest that the appraiser faces a practically impossible task if the appraisal is to be done correctly. Luckily, Principle 2 below will show that we can usually escape this conundrum relatively easily. We shall return to the secondary profits question a shortly, therefore.

Principle 2: Judgements about the nature and general state of the economy in question provide the key to making choices about which kind of secondary impacts, if any, should be included as cost and/or benefit flows in an economic appraisal.

To illustrate what this rather ambiguous notion might entail, let us think about the estimation of changes in secondary profits as a consequence of the project. I continue to follow relatively closely what Tresch has to say on this matter. He argues that to a good approximation econometric evidence from a wide variety of sources suggests that an economy's aggregate production function will be characterised by constant returns to scale (CRS). If we accept this proposition, then a well-known result from economic theory tells us that in the long run all pure profits in the economy will be eliminated; supernormal profits will be transitory (short-run) in duration. He then argues that the short run profit effects, while undoubtedly being both large and significant at the individual firm or industry level, will tend towards a zero sum in aggregate. This is because positive effects on profits (mainly arising in industries whose activities are complementary to for the project's activities) are likely to be exactly balanced by negative effects on profits (mainly arising in industries whose activities are substitutes for the project's activities). Unless we have specific knowledge that would lead us to reject this zero sum aggregate impact, then the safest thing is to ignore secondary profit impacts in an economic appraisal.

We have argued here that impacts on secondary profits should not be included in an economic appraisal. What about profit effects directly associated with the project? they SHOULD be included. But note that it is possible that they will already have been factored into the CBA benefit calculations. One way of thinking about benefits is that they are increases in consumer and producer surpluses. Changes in profits are, of course, equal to changes in producer surplus. If we have already included in our benefit estimates increases in producer surpluses in firms directly involved in the project, then including the profits as well would constitute pure double counting. If we have not, they should be included.

How does including producer surpluses square with Table 11.7 which mentions only utility changes? This is easy. Firms are owned by households, and so business profits are (ultimately) paid to individuals and so affect utility functions. Clearly there is no inconsistency.

Now let us examine what position one should take regarding the 'multiplier' impacts of a HEP project. I shall again quote at length here (from Tresch, 1981, page 559) as it raises immediately the central issues:

"Something must be wrong here, for it appears that any project placed in an underdeveloped region of a country must be worthwhile. The crux of the problem is that Keynesian multiplier analysis is simply irrelevant to fundamental goal of cost-benefit analysis, determining the best use of a society's scarce resources. Cost-benefit analysis begins with a presumption of full-employment so that, strictly speaking, it is concerned with the maximum expansion of society's production-possibilities frontier. Keynesian multiplier analysis, on the other hand, is concerned with moving an unemployed economy to its production-possibilities frontier. Hence, if the full-employment assumption is retained, the multiplied increases in income associated with any one project will be exactly offset by multiplied decreases in income in other regions of the economy which lose resources to the hydroelectric project.

The kernel of truth here is that market economies are seldom fully employed and that cost-benefit analysis should be adjusted to account for unemployment. If a project creates a net gain in employment for the economy as a whole, this is a short-run benefit that can legitimately be included. But is it prudent to include these gains? .... Unless one can argue that convincingly that some particular project will have unusually strong multiplier effects, there seems to be little point in attaching a multiplier analysis to a cost-benefit study."

It is worth looking at some of these points a little further.

Cost-benefit analysis begins with a presumption of full-employment

In general, the most safe assumption to make about the state of the economy is that it is characterised by full employment. Thus, new projects will divert resources away from other activities, and will incur opportunity costs given by the return available to those resources in their best alternative use.

Taking full employment as the default position in an economic appraisal also helps to safeguard against optimism bias (see a later page) and creates some immunity to the pleas made by special interests that have vested interests in the promotion of a particular project.

Having a full employment presumption as the default position does not mean that it must be accepted in all circumstances. But it puts the onus on the appraisal team to give reasoned arguments, and evidence-based support, for proceeding on any other basis.

If the full-employment assumption is retained, the multiplied increases in income associated with any one project will be exactly offset by multiplied decreases in income in other regions of the economy

An extension of the previous opportunity cost argument, as Tresch points out, is that indirect incomes generated through multiplier impacts will be offset by losses in incomes elsewhere. To include the generated indirect incomes without deducting the lost incomes elsewhere would constitute an error equivalent in effect to that of double counting.

Market economies are seldom fully employed and cost-benefit analysis should be adjusted to account for unemployment.

Even if an appraisal team cannot be confident that the economy as a whole is one characterised by general involuntary unemployment, it may sometimes be the case that the direct impacts of a project may be specifically designed to make use of unemployed or under-employed members of the workforce. Where one can be confident about this, then an appropriate way of dealing with this in an EA is to use an appropriate shadow wage rate for those individuals who come into employment from unemployment as a result of the project. The appropriate shadow price of labour in these circumstances is likely to be well below market wage rates for the kind of work in question; it will only be zero, though, if the individuals in question did not have a positive marginal value on their leisure time. [If that marginal valuation were actually negative, then the shadow wage rate would change sign.]

None of this alters the fact that even if the particular workers taken on in the project were known to be unemployed (or unemployable) without the project, we might still wish to work with the presumption of full employment overall in the economy (in which case indirect effects should not be included). In the words of Tresch, 'unless one can argue that convincingly that some particular project will have unusually strong multiplier effects, there seems to be little point in attaching a multiplier analysis to a cost-benefit study.'

Indirect impacts are short-run only?

Most of the theoretical literature regarding cost-benefit analysis was written by authors working in and thinking in the main about affluent, developed economies. In those settings, it was common for economists to think of involuntary unemployment as being a short run, business cycle, phenomenon, associated with demand or supply side shocks to the economy. Economies moved in the medium (and long) run towards states in which labour and product markets cleared. It is this background which underpins the view that multiplier-type effects are, at most, short run and so transitory in duration.

Whether this is an appropriate way of thinking about low-income developing economies is at best a moot point. Many development economists take the view that such economies are capital scarce but labour rich, with a long term structural tendency towards generalised excess supply of labour. If the appraiser is satisfied that the economy being studied is one in which this is the case, then the standard position that indirect and induced impacts should not be included loses its validity.

E.4 Some additional points regarding secondary impacts

We have said nothing explicitly to this point about projects that have a specifically regional, or regional development, focus. However, the points that have been made above apply here too.

In affluent developed economies (such as the UK, the other EU economies, and the USA), so-called "regional policy" has long been an important part of social, political and economy policy. The emphasis here has been on trying to use preferential policies to bring the economic performance of more slowly growing regions within a national economy closer to that of its more prosperous or faster growing regions. One aspect of this has been policies or programmes that seek to boost employment levels (usually by means of labour or capital subsidies, either applied generally throughout a particular region or applied selectively on a discretionary basis).

Appraising and evaluating (ex ante or ex post) such programmes involves dealing with questions about additionality. In simple terms, labour and/or capital subsidies or other forms of special assistance can create jobs, but the net impact on employment and output will be less than the gross impact, and in many cases substantially less. In extreme cases, the net impact may be zero.

The UK Treasury (2003) Green Book discusses issues about additionality and the gross versus net impact question at some length in its Appendix 1. The "Additionality" button at the foot of this page takes you to an extract from the Green Book that deals with these issues. If you have not looked at this material before, you will find that reading these couple of pages is very worthwhile.

DFID activity does not of course take place in affluent developed economies. But many lower income developing countries have regions that are relatively under-developed or which have particularly high levels of social and economic deprivation. Projects or programmes might be developed for dealing with such imbalances. Once again, the question arises about which secondary impacts should be included in an economic appraisal. The answer to this question is the same as that we have given already. One must include and account for secondary impacts to the extent that we have reason to believe that the beneficial secondary impacts will be less than fully offset by corresponding adverse secondary impacts elsewhere. Our judgement will be driven in the main by information that is available regarding the short and medium term state of the economy. It is very important, though, that the appraiser does not fall into an extreme form of 'optimism bias' by including only positive multiplier-type impacts and not giving proper attention to offsetting impacts elsewhere.

E.5 Large Projects: Some Further Considerations

In some early texts, it was argued that the technique of Cost Benefit Analysis (CBA), as generally used, dealt only with the analysis of projects that were marginal with respect to the economy as a whole. If this position were to be accepted, then policies that were intended to change the nature of the economy would not be amenable to CBA (and perhaps not to what is commonly taken to be economic appraisal). Thus, for example, a policy decision to abandon the market system in favour of command and control, is not marginal, and so would not be amenable to analysis using CBA. However, a policy decision to introduce a new form of taxation would be marginal. And an investment project, such as a new nuclear power plant or a new airport, could be large in absolute terms, but would nonetheless be a small part of total investment, and hence marginal. Such schemes therefore would be amenable to CBA.

However, the distinction between marginal and non-marginal projects or policies is not always easy to make in practice, and some economists would argue that any policy or project that either changes relative prices and/or alters consumer and/or producer surpluses is non-marginal. If so, accepting a proposition that CBA (and economic appraisal more generally) is only applicable to marginal changes would exclude much of what it is that we would wish to evaluate. This is especially true for project work in developing economies, where much of what is of interest is about harnessing major structural changes in the economy.

Nowadays, few economists would accept that CBA is necessarily restricted to dealing with projects that are marginal with respect to the economy. There are, of course, some important implications of adopting this 'wider' perspective.

First, it implies that the analyst can no longer rely exclusively on partial equilibrium modelling techniques, as these presume that impacts in the rest of the system are trivial. But where projects are large relative to the rest of the economy, and/or where projects have significant supply-side impacts, then relative prices will change throughout the economy, inducing a variety of substitution effects. In principle, to get a proper handle on these impacts the analyst must make use of system-wide general equilibrium modelling techniques. (We discuss CGE models in a page of voluntary additional reading elsewhere in this Topic). Alternatively of course, one might invoke the arguments used above that prior knowledge about the state of the economy allows us to make rough estimates as to what these general equilibrium impacts would be (or to argue that they will in aggregate cancel each other out to a reasonable approximation.)

Second, when it comes to evaluating impacts, marginal valuations may no longer be sufficient. Projects may produce large changes in output or demand, so that valuations can no longer make use of just marginal costs or benefits. In these cases, we will need information about demand and cost schedules so that changes in consumer and/or producer surpluses can be estimated.

Third, and possibly most difficult, are issues relating to external economies of scale. For sufficiently large projects, something akin to an agglomeration or cluster effects can arise, substantially altering a regions economic competitiveness. This might be true, for example, in the case of major transport of electrification infrastructure projects. The estimation of the values arising from strategically important transport infrastructural improvements have proven very difficult to deal with satisfactorily because of their multi-faceted nature, and because of the uncertainties that arise as a result of potential relocations of economic activity that such projects might induce.

F. Public Sector Bias?

Some practioners of economic appraisal often feel uncomfortable when faced with suggestions that public sector projects face an easier set of hurdles than do private sector investment projects. Such suggestions then extend to the notion that there is an inherent 'pro public sector bias' as a result of the way that economic appraisal is carried out. Is there any truth to such claims?

To answer this question, let us look at the specific sources of bias that some commentators have in mind. There are two principal claims here:

1. Government projects are evaluated at a far lower discount rate than are private sector projects (and so have a lower hurdle to overcome).
2. In evaluating decreasing cost projects (i.e. projects that involve activities for which the long run average cost curve is declining over all relevant output ranges), public sector appraisal techniques allow for projects to be accepted even if they are not financially profitable whereas that option is not feasible for the private sector.

We now investigate each of these two claims.

1. Government projects are evaluated at a far lower discount rate than are private sector projects (and so have a lower hurdle to overcome).

As far as discounting is concerned:

- the relevant rate of discount for public sector projects is the social discount rate.
- the relevant rate of discount for private firms is given by either by the cost of capital (which will incorporate a risk premium for what the market takes to be the riskiness of the activity in question) or by the gross-of-tax rate of return (which is, in effect, the marginal rate of transformation, MRT, between the present and the future).

Let us label the social discount rate as r_s and the private discount rate as r_p. In most circumstances, r_s will be less than r_p, and sometimes very much less. But as we shall see in a moment, this is what **should be** the case; it is not a bias as such. Before we think about that matter, a more fundamental point needs to be made. Correct implementation of economic appraisal (as we saw right at the start of this course) requires that there is a rationale for the intervention in question. A valid rationale for intervention will typically have either an efficiency-enhancing justification or an equity-enhancing justification (or both). Hence, public sector interventions are being done in circumstances where

- either private sector involvement would not be forthcoming, or would not be forthcoming to the extent or that is desired (the provision of public goods, such as public health, would be an example), or
- market -led outcomes are deemed to be less than optimal in terms of outcome on some ethical or distributional criterion.

The point here is that it is inappropriate to think in terms of there being a set of exogenously-given projects that could or would be done equally well either by the private sector or by the public sector. Put another way, we are **not** dealing with cases where all else is equal and the only difference is whether the project is done by the private sector at r_p or by the public sector at r_s.

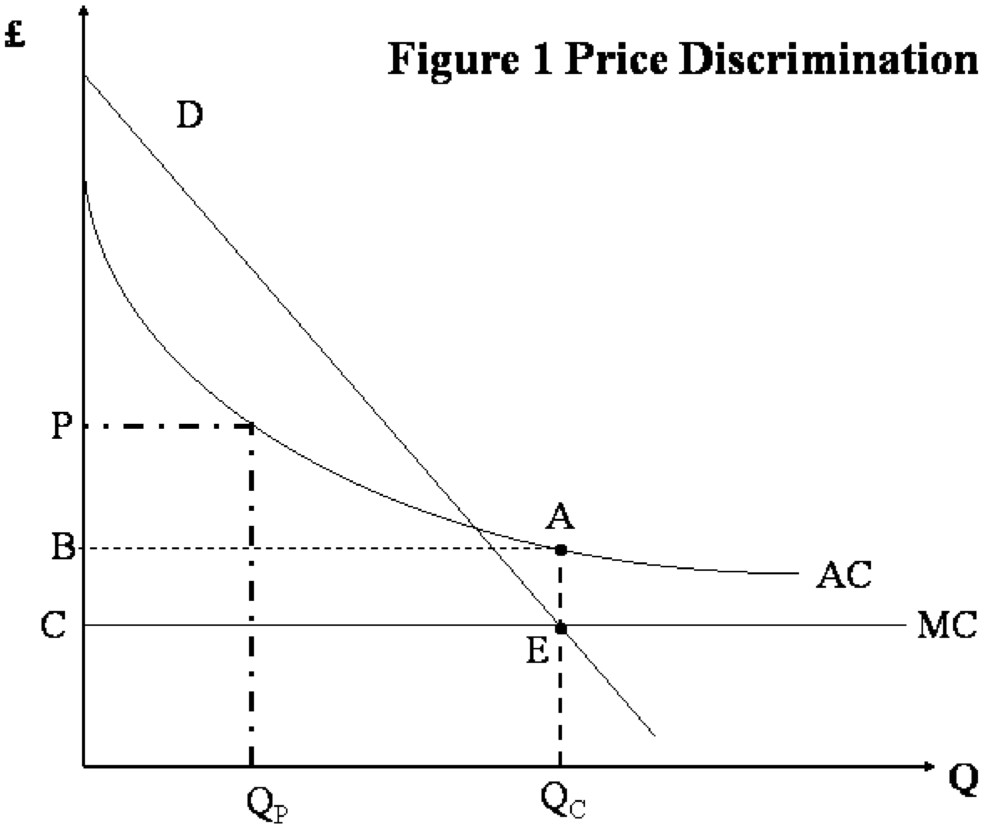
Once one accepts that there is a socially-valid rationale for an intervention, then what should the appropriate discount rate be for such projects. The answer is, **by definition**, the social discount rate. You have already seen that the choice of a precise number for the social discount rate is by no means an easy matter. But there are many reasons why any sensible choice of discount rate will yield a figure that is different from (and is often lower than) the private sector discount rate. These reasons include:

- the government has a greater ability to spread risks than does the private sector; this is a real cost advantage held by the government sector, and so risk-premia can be far lower than would be the case with private projects;
- government and private projects are often fundamentally different in terms of their time profiles of costs and benefits, and they can also be very different in terms of patterns of private consumption and investment; given such differences, there can be no general presumption that r_s and r_p should be identical;
- intertemporal externalities should be taken account of in government project appraisal, and one way of doing so may be through the value at which the r_s is set.

2. Decreasing cost projects

Recall that the point sometimes made here is that for projects for which the long run average cost curve is declining over all relevant output ranges, public sector appraisal techniques allow for projects to be accepted even if they are not financially profitable whereas that option is not feasible for the private sector.

An obvious case in point is provision by the state of natural monopolies for utility industries with increasing returns to scale arising from indivisible network costs. We can think through this case with the help of Figure 1 below. This represents a natural monopolist facing demand curve D with constant marginal cost C and average cost AC. First-best efficient pricing requires that price be set equal to marginal cost. If done, this would produce output Qc at a price equal to C and the firm would make a loss of CBAE.



No private firm could (without subsidy) price at marginal cost, as it would entail making negative profits. Of course, the government might allow firms to set price equal to average cost, to just break even, but this solution - while being second-best - is still inefficient relative to the first-best marginal cost pricing outcome.

Suppose that we regard public ownership along with marginal cost pricing as a project for which an economic appraisal is being carried out. Would the project be accepted? If we ignore the possibility that public ownership is intrinsically more inefficient than private ownership (which if may, of course, be) then such a project would be accepted. To see why, compare total benefits with total costs. Total costs are given by average costs multiplied by output, that is the area formed by the rectangle that arises when output (the distance BA) is multiplied by average cost (the distance AQ_C). Total benefits are equal to the area under the demand curve from zero output up to output Q_C. This area contains, of course, the consumer surplus that is being generated when output is at $Q = Q_C$. Equivalently, one might argue that the financial loss is more than compensated for by the consumers' surplus gained (consumers could compensate the producer and remain better off). [See *Figure 1a* in the PowerPoint file *note3a.ppt*.]

The point here is NOT to argue that state ownership of public utilities is a good thing, nor to argue that there are not other ways in which efficient outcomes could be achieved even if there were private ownership (which there are). Rather, it is to make the point that in carrying out an economic appraisal, it is correct and proper to take account of consumer (and producer) surplus. The fact that private businesses cannot do so because it is not revenue that accrues to them is simply irrelevant when it comes to decisions about the efficient allocation of resources.

Further Notes to Accompany Page E: Green Book Guidance (Annex 1 of 2003 Green Book, on Rationale for Intervention)

Additionality

12. The success of government intervention in terms of increasing output or employment in a given target area is usually assessed in terms of its additionality. This is its net, rather than its gross, impact after making allowances for what would have happened in the absence of the intervention. Additionality can also be referred to as a supply side impact, which operates by altering the productive capacity of the economy. This can occur either because of a change in the size of the workforce or a change in the productivity of the workforce. Examples of interventions that promote supply-side benefits include improving the working of markets and economic institutions, strengthening capabilities, and facilitating greater participation in the workforce. The extent to which a proposal may produce a supply side benefit is an important component of an appraisal.

13. If there are no grounds for expecting a proposal to have a supply side effect, any increase in government expenditure would result in a matching decrease in private expenditure, (known as crowding out. If, however, the supply-side impact of a proposal is expected to be positive, the net additional impact on economic welfare will need to be measured. This may consist of additional employment or output, and constitutes a real net benefit which the appraisal should take into account.

14. Estimating this type of additionality will normally require an analysis of the product, labour, and in some cases, capital markets affected by the intervention. For example, when assessing the level of displacement of an employment creation programme or the impact of recruitment and redundancy decisions on a particular local area, it is necessary to examine the characteristics of the jobs created, or protected, in relation to the characteristics of the local labour market. They must then be compared with similar jobs in other local areas that are not subject to the policy. Such a comparison establishes the do nothing case: what would have happened if the intervention had not gone ahead.

15. In some cases, the best source of information for assessing additionality may be from those who clearly have an interest in the outcome of the decision. In these circumstances, the information and forecasts should be confirmed by an independent source. For example, the implied growth in demand for services might be compared to other forecasts for the same region, and contrasted with past performance. Sensitivity analysis should also be carried out, using alternative values for the key variables.

16. After developing the do nothing case, the next step is to assess the net impact or benefit of these different options. This net benefit is the additionality of the option. Additionality must, however, be calculated with consideration of 'leakage', 'deadweight', 'displacement' and 'substitution' effects. These are explained below.

- Leakage effects benefit those outside of the spatial area or group which the intervention is intended to benefit.

- Deadweight refers to outcomes which would have occurred without intervention. Its scale can be estimated by assessing what would have happened in the 'do minimum' case, ensuring that due allowance is made for the other impacts which impact on net additionality.

- Displacement and substitution impacts are closely related. They measure the extent to which the benefits of a project are offset by reductions of output or employment elsewhere.

17. For example, a project may attract scarce skills, or investment, which would otherwise have gone to other parts of the country; or, if the policy involves support for local businesses, these may compete for resources and / or market share with non-assisted businesses.

18. The appropriate area for analysis of displacement effects will depend on the type of project. In the case of employment displacement, the area considered should usually approximate the local labour market. [Detailed guidance on methodologies for assessing displacement effects is available from the DTI Central Evaluation Team web site at http://www.dti.gov.uk.The recent DTI/ SBS evaluation of 'Smart', available on the same web site, provides an applied example. Also useful is research undertaken for DTI by the University of Durham (http://www.dur.ac.uk) and DWP s Travel to Work Areas.]

19. The effect on net employment and net output is likely to be much smaller than the direct employment and output effects of the project. Evidence should support the assessment of the scale and importance of any net employment and net output benefits, taking account of multiplier effects. A multiplier measures the further economic activity, (whether output or jobs), resulting from the creation of additional local economic activity. Where it is considered appropriate to calculate multipliers, guidance is available from English Partnerships and the Regional Development Agencies. [For example, see 'Additionality: A Full Guide' (English Partnerships, 2001)]

> The net benefit of an intervention equals the gross benefits less the benefits that would have occurred in the absence of intervention (the deadweight less the negative impacts elsewhere (including displacement of activity), plus multiplier effects.

20. If there is no improvement in national economic efficiency, local employment and output effects, net of any local displacement effects, may be considered in parts of the appraisal where the project has a strong distributional rationale. For example, a policy may aim to reduce the rate of unemployment in a particular deprived area, as opposed to reducing the rate of unemployment overall.

21. Where potentially large changes to employment, (either as a result of employment creation, protection or redundancy) are concerned, assessment will normally require a thorough analysis of the local labour market. This should cover the age, skills and experience of those whose jobs are at stake, and how these compare with the characteristics of the unemployed and those who have recently found employment. The

analysis might also assess the likelihood of new investment in the region in the event that these job losses occurred.

Computable General Equilibrium (CGE) Analysis

Suggested Further Reading :

The notes you will find on these pages give a simple introduction to CGE modelling. Should you wish to study this topic further, four references are given below for further reading. The Greenaway et al reference is a particularly good starting point for understanding CGE as it was written for policy makers, and so assumes nothing (in terms of technical knowledge.)

1. Greenaway, D., Leybourne, S.J., Reed, G.V. and J. Whalley (1993), Applied General Equilibrium Modelling: Applications, Limitations and Future Developments, HMSO, London .
2. Pyatt, G. and J.I. Round (1985) eds. ***Social Accounting Matrices: A Basis for Planning*,** The World Bank, Washington , D.C. , U.S.A.
3. Reinert, K. A. and D. W. Roland-Holst (1997) "Social Accounting Matrices", J. F. Francois and K. A. Reinert (eds), *Applied Methods for Trade Policy Analysis: A Handbook*, Cambridge University Press, Cambridge.
4. Devarajan, S., Go, D.S., Lewis, J.D., Robinson, S. and P. Sinko (1997), "Simple General Equilibrium Modelling" in J.F. Francois and K. A. Reinert eds. Applied Methods for Trade Policy Analysis, Cambridge University Press, Cambridge.

A Basic Introduction To General Equilibrium Modelling

Applied or computable general equilibrium (CGE) analysis involves simulating numerically the general equilibrium structure of the economy. The use of the term 'general equilibrium', and the framework upon which applied general equilibrium models are built, corresponds to the theoretical Walrasian general equilibrium system.

Definitions:

Mankiw (2003, Macroeconomics, 5th edition, page 530) defines **general equilibrium** as "The simultaneous equilibrium of all the markets in the economy".

Walrasian general equilibrium refers to the general equilibrium model with production introduced in Léon Walras's *Elements of Pure Economics* (1874). The basics of the model (as first outlined by Walras) are the following: individuals are endowed with factors and demand produced goods; firms demand factors and produce goods. General equilibrium is defined as a set of factor prices and output prices such that the relevant quantities demanded and supplied in each market are equal to each other, i.e. both output and factor markets clear.

Source: "History of Economic Thought" web page at http://cepa.newschool.edu/het/essays/get/getcont.htm

See also "General Equilibrium" (and related articles) in the New Palgrave Online Dictionary.

Essential feature

The **essential feature** of this Walrasian general equilibrium system is that supply equals demand **in all markets** at a set of relative prices that can be identified. Usually, but not exclusively, the basic underlying theoretical structure in CGE models corresponds to that developed and refined by Arrow, Debreu and Hahn (see Arrow and Hahn, 1971).

The conventional Walrasian model assumes universal perfect competition. However, in practice general equilibrium models are **not** restricted to conditions of universal perfect competition (as you will see below). In terms of essential assumptions, all that is needed is that an equilibrium for the economy exists and that it is unique. (Greenaway *et al* , 1993, Shoven and Whalley, 1984)

Greenaway *et al* (1993) give a succinct outline of the key issues in laying out the basic structure of a general equilibrium model, which is the first stage in any CGE analysis. The four key issues are:

- ***Dimensionality :*** the level of sectoral disaggregation of total economic activity (i.e. the number of products/production sectors and factors of production)
- ***General specification of key relationships***(including functional form) - supply and demand equations (including the interdependencies/interactions between sectors)
- ***Collection of benchmark data*** : to model the benchmark case/initial equilibrium
- ***Calibration of the model's parameters to that data set*** : while key parameter values will be pre-specified, calibration involves choosing the remaining parameter values so that the model can reproduce the data set as an equilibrium solution (Shoven and Whalley, 1984)

Once the structure is in place the model is solved for general equilibrium. Any general equilibrium model with price endogeneity and many sectors will generally be impossible to solve algebraically. However the existence of sophisticated algorithms and powerful computing facilities mean that such systems can be solved computationally, data availability being the principal constraint in the presence of appropriate hardware and software. (Greenaway *et al*, 1993)

After the model structure has been determined and the model has been solved for general equilibrium, a number of exercises can be carried out, including:

- equilibrium relationships and interdependencies can be traced and examined
- counterfactual equilibria can be computed for exogenous changes
- questions of policy evaluation, including distributional effects, can be addressed.

The CGE framework is a very flexible one, within which many different models can be constructed, each with different variations on the basic structure, as will become clear later. So it is useful to go through these issues for a particular type of application

Input-output modelling is an example of general equilibrium modelling. Indeed, input-output (IO) modelling is the most familiar and widely employed general equilibrium framework. It is, however, very restrictive and because of its assumptions it can be regarded as one of two limiting cases, in which supply is entirely passive and the model is demand-driven.

The other limiting case is the classical model of perfectly competitive markets, where relative prices adjust so that all markets clear. One can also think about the perfectly competitive market model as one in which demand is entirely passive and the model is supply-driven.

So we have two special cases of general equilibrium system that can be envisaged as lying at opposite poles of a continuum. These two special cases are encompassed by the more flexible computable general equilibrium (CGE) modelling framework, which include these two extremes as special cases but can also accommodate all intermediate positions. (See Harrigan and McGregor, 1988, and McGregor *et al* , 1996 for further details of this assertion).

Additional References:

Arrow, K.J. and F. Hahn (1971) *General Competitive Analysis,* San Francisco, Holden-Day.

Greenaway, D., S.J. Leyborne, G.V. Reed and J. Whalley (1993) *Applied General Equilibrium Modelling: Applications, Limitations and Future Development,* HMSO.

Harrigan, F. and P.G. McGregor (1988) "Price and Quantity Interaction in Regional Economic Models: The Importance of 'Openness' and 'Closures'" in F. Harrigan and P.G. McGregor eds. *Recent Advances in Regional Economic Modelling,* Pion Limited.

McGregor, P.G., J.K. Swales and Y.P. Yin (1996) "A Long-Run Interpretation of Regional Input-Output Analysis", *Journal of Regional Science,* Vol.36, No.3, pp.479-501.

Shoven, J.B. and J. Whalley (1984) "Applied General-Equilibrium Models of Taxation and International Trade: An Introduction and Survey", *Journal of Economic Literature,* Vol.52, pp.1007-1051.

The relationship between Input-Output Modelling and CGE Modelling: Input-Output As A Limiting Case Of General CGE Modelling

We noted above that IO is the most familiar, widely employed, but restrictive, method of general equilibrium modelling. The assumptions of fixed technical coefficients in production and a perfectly elastic supply side mean that IO is normally associated with short-run modelling in an imperfectly competitive, excess capacity setting.

However IO can be, and has been, interpreted as a special, limiting case of a neo-classical general equilibrium system, with its results representing long-run equilibria where supply constraints do not bind.

Let us begin by using a classification of the 'calibration procedures and model use' identified by Shoven and Whalley, by means of which they give a set of criteria that are useful to describe and define the kind of CGE model one is using. [The reference is Shoven and Whalley, 1984, page 1019, Fig.1: "Flow chart outlining calibration procedures and model use in typical applied general equilibrium model" , which is also reproduced in Greenaway *et al* (1993), page 23, Figure 2.1 .]

If we think about IO analysis as one special case of CGE modelling, let us use Shoven and Whalley's criteria to characterise IO analysis:

1. Dimensionality

What and how many sectors from the IO accounts should be identified in the CGE model? This is largely a matter of the degree of aggregation/disaggregation of production sectors and components of final demand that we choose to use.

2. Data collection.

Data on sectoral demand and supply, factor use and rewards, total GDP etc are collected (using primary and/or secondary sources) to build the transactions table - which is essentially an expanded (disaggregated) extension of the national accounts where inter-sectoral flows of goods and services are explicitly included. You have seen, for example, that the Scottish Executive produces a 128 x 128 industry by industry IO table. This would form the benchmark equilibrium data set for a CGE model of the Scottish economy, for example.

Since data are often drawn from a variety of sources, including surveys, the initial table is likely to require adjustments for consistency: in IO tables the value of each sector's total input must balance with its output. For the table as a whole, total intermediate sales must equal total intermediate purchases, and total final demand must equal total primary input. These are accounting conventions (income = expenditure) which must hold for consistency.

3. Choice of functional form

To move from accounting (transactions/IO tables) to IO **modelling**, the conventional basic IO assumption is that:

(1) $X_{ij} = a_{ij}X_j$

where a_{ij} is a *constant*. What this tells us is that IO assumes fixed technical coefficients in production: Leontief technology.

Thus, all IO relationships are of a fixed linear form, and are not subject to any supply constraints.

4. Calibration to benchmark equilibrium

Setting up a CGE model require that the model has first been calibrated, and that the model can recreate the base year data. What does this mean and what does it entail?

We have to begin by choosing one particular year as our base year. The nature of CGE modelling requires that we assume that the economy was in long-run equilibrium in this base year. (Not surprisingly, this is a rather controversial assumption).

A CGE model will consist of a set of "structural equations" that specify, among other things, demand and supply in the various markets we have identified in the model. These structural equations will relate the values of variables to one another. Parameters in these equations will determine the quantitative nature of these relationships.

Parameters and variables: what are these?

Students often confuse parameters and variables. to explain the difference, think about a very simple model in which aggregate consumers' expenditure (C) is a linear function of aggregate consumers' disposable income (Y). This 'structural equation' can be written in the form:

$C = a + bY$

Which of these are variables and which are parameters?

Variables

In this equation, there are two variables, C and Y. In a general equilibrium model, each of these is likely to be an endogenous variable, the value of which is determined by the simultaneous solution of all the structural equations in the model. In terms of earlier language, this also means they are the values which C and Y will take when all markets are simultaneously in equilibrium.

The researcher will typically have a benchmark data set (the SAM, to be described and explained later) which gives us the values taken by C and Y (and all others variables in the CGE model too) in the base year. So C and Y are known in the base year. You should also be able to see now why it is necessary that we assume that the economy was in general equilibrium (supply equals demand in all markets) in the base year.

Parameters

So C and Y are variables. What, then, are the parameters? In this equation, there are two parameters, a and b. Parameters govern the magnitudes in the relationship between the variables. For example, if a = 100 and b = 0.5, then the equation becomes

$C = 100 + 0.5Y$

and so for any particular value of Y there will be a unique value that C must take. Clearly, different values of a or b, or both, would lead to different relationships between C and Y, even though the form of the relationship would still be the same qualitatively.

Although some parameter values will be known in advance, many are unknown numbers and so values must be found for them somehow. That may involve **estimation** using the regression and statistical inference techniques covered in the DHA class, for example), or it may be done by calibration, as we shall explain below.

For any CGE modelling exercise, we use prior information (where it is available) to specify the numerical values of some parameters. Some other parameter values will be unknown but can be estimated statistically; where this can be done reliably, those estimates will be used as proxies for the true but unknown numerical values of the parameters.

But inevitably, not all parameters can be assigned values in these two ways. Those that cannot must be assigned values by means of **calibration to the base year**. Calibration to the base year equilibrium involves specifying the value of some parameters (using the two methods described above) and then running the model so that it recreates the base year, **solving for all unknown parameters**. Put another way, we select or 'calibrate' a set of parameter values which have the property that - for this particular set of parameter values - the model, when run with no changes in any variables from their baseline values, will generate results that are identical too, and so consistent with, the baseline data.

In the case where the CGE model we are investigating takes the special case of being an IO model, the parameters will largely consist of technical coefficients that can be deduced directly from our base year data. The calibration step straightforward in this case. Suppose X_j is the output of the manufacturing sector, j, and X_{ij} records purchases from the utilities sector, i, to the manufacturing sector, j. This means that we are assuming that utilities purchases (e.g. electricity) are a constant fraction, a_{ij}, of the value of manufacturing output. It is these coefficients that are the parameters of the IO model special case of CGE. That is, if manufacturing output doubles, so will its required input from the utilities sector. Thus, all IO relationships are of a fixed linear form, and are not subject to any supply constraints. From this a table of intermediate input coefficients is constructed, from which the basic IO model is built.

5. Analysis using model.

With the model calibrated for the benchmark equilibrium, in the case of IO, exogenous demand changes (e.g. policy) can be specified, counterfactual equilibria can be computed, and policy appraisal carried out based on comparison between the counterfactual and benchmark equilibria.

You have seen that IO systems are conventionally *demand driven*, with effects typically stemming from changes in the final demand sector. However, an important deficiency of conventional IO is that, given the demand-driven focus, the model is silent on prices.

Given the assumption of no supply constraints, the implicit assumption is that there are similarly no price changes in response to changes in final demand. It is for this reason that we can describe an IO form of CGE model as a **Keynesian** model. Recall from standard macro theory that a Keynesian model is one in which changes in exogenous demands (such as additional government spending or upward shifts in capital expenditure) lead to **quantity changes in the model, but no change in prices**. So changes in output, GDP, income and employment are all demand driven, and are all "real" or volume changes, as prices are unaffected by demand.

There are no consumption or income effects **induced by price changes** (though IO **can** incorporate *non-price-induced* consumption and income effects in Type II IO multipliers).

Also, because input coefficients are constant, there can be no input substitution effects and, with its assumption of a passive supply-side, evaluation of supply-side policies are necessarily deficient.

This is a particularly important point when regional and environmental policies are in question, since these most commonly take the form of supply-side policies.

To attempt evaluation, taking account of supply-side factors, these effects would have to be worked out 'off-line' (outside the model), on an *ad hoc* basis, then the modeller needs to go back and change the model itself.

In summary, IO can be regarded as a simple general equilibrium system, with fixed coefficient technologies, an absence of capacity constraints, and an infinitely elastic supply of primary factors.

The strengths of the approach are the ease of implementation and potentially high levels of sectoral disaggregation, allowing interactions and interdependencies to be easily traced and identified.

The conventional interpretation of IO is that it applies to short-run (impact period) in an imperfectly competitive setting where there are conditions of general excess factor supply. However, there is an alternative interpretation of IO: that IO may replicate the long-run equilibrium of a neo-classical model.

McGregor *et al* (1996) demonstrate that in a wide range of regional economic models, IO accurately describes the long-run response of activity to a permanent change in aggregate final demand. They explain that many regional models do not operate as IO systems in the short-run, but may do so in the long run.

This is especially the case in a regional context, where perfect inter-regional factor mobility might be approximated in the medium to long run.

Here, let's start with identifying the base year data for a more flexible CGE framework.

References:

Greenaway, D., S.J. Leyborne, G.V. Reed and J. Whalley (1993) ***Applied General Equilibrium Modelling: Applications, Limitations and Future Development***, HMSO.

McGregor, P.G., J.K. Swales and Y.P. Yin (1996) "A Long-Run Interpretation of Regional Input-Output Analysis", ***Journal of Regional Science***, Vol.36, No.3, pp.479-501.

Shoven, J.B. and J. Whalley (1984) "Applied General-Equilibrium Models of Taxation and International Trade: An Introduction and Survey", ***Journal of Economic Literature***, Vol.52, pp.1007-1051.

THE SOCIAL ACCOUNTING MATRIX (SAM)

When reading this page, please try to get the intuition and get the overall picture. The details are not important - you can always pick them up again later should you ever wish to do work in this area.

A Social Accounting Matrix plays an equivalent role in CGE analysis to that played by a transactions table in an IO model. In essence, each is the database for the model. The key difference is that a SAM contains a much wider variety and depth of information. We begin here by examining what is - and is not - in an IO transaction table, and the explain the content and structure of a typical SAM.

1. IO Tables: what they tell us and do not tell us

In either national or regional accounting terms:

- An Input-Output (IO) table tells us how income is generated within the target economy during the time period in question
- That is, it tells us about the composition of the economy's gross domestic product (GDP) by describing the structure of production.

What IO doesn't tell us:

- How this income is **distributed** among the main transactors who engage in economic activity in the target economy?
- How much of the income generated within the local economy actually accrues to the people who live within its boundaries?
- How much flows to external transactors (for example, in the form of repatriated profits)?
- How much income residents of the local economy earn as a result of income generation that takes place elsewhere?
- Taking the three previous points together, IO fails to tell us anything about the distinction between Gross Domestic Product (GDP) and Gross National Product (GNP).

As well as production, the IO table tells us about the consumption/expenditure patterns of the different types of transactor.

However, it does not tell us much about the income side of the equation.

- In the case of Households, the IO table only tells us about wage income
- Not likely to be sufficient to finance total household expenditure.
- Where does the additional income required to fund consumption come from?
- Is it generated from some local source not reported in the IO, or abroad? How much is saved and invested?

If we want to properly model and analyse behaviour in the economy, we must incorporate more comprehensive information on income and expenditure flows in the database that provides the statement of initial equilibrium conditions. **The framework that fulfils this requirement is a social accounting matrix (SAM).**

Later in these notes, we give details on the schematic SAM used in the AMOS modelling project, a major CGE modelling project that has been running in the University of Strathclyde over the last 20 years. The actual data for the 3 sector model are also shown later.

The following image provides a schematic, generic representation of the form and components of a SAM. We shall refer to this graphic several times below. Below the schema is an explanation of each of the elements in the SAM.

Figure 1: A schematic Social Accounting Matrix (SAM)

Expenditure by / Income to	Production Activities (the I production sectors)	Institutions H C G CF E	Factors of Production Labour (L) Capital (K)
Production Activities	**T**	**U**	
Institutions: H C G CF E	**V**	***W***	***X***
Factors of Production: L (Value Added) K	**Y**		

Explanation of each of the elements in the schematic SAM:

The bold capital letters in each of the blocked areas represent a sub-matrix of cells showing, along the rows, the value of the income flow to the transactor identified down the left-hand side of the SAM from the transactor identified along the top of the SAM. Where no letter is shown in a block, this means that no interaction takes place: e.g. none of the production sectors receive income from either of the factors of production. Where the letter is shown in standard bold format, this means that data can be taken directly from the IO table; where the letter is shown in bold italic, this means additional data are required for construction of the SAM.

- Production sectors are identified by the subscript i (where i = 1,..,I)
- Institutions/aggregate transactor groups by a (where a = 1,..A)
- Factors of production by b (where b = 1,□,B).

Therefore

T is an IxI matrix of intersectoral transactions between the I production sectors of the economy	Sub-matrix **T** is identical to the upper left quadrant of the IO table
U is an IxA matrix of final demand expenditures by the A institutional transactors identified above on the outputs	The entries for sub-matrix **U** are given by the final demand block (upper right quadrant) of the IO table (the aggregate transactor 'Corporate' is not a final

of the I local production sectors.	demander so all the column entries for C in matrix **U** are equal to zero).
V is an AxI matrix of income flows from the I production sectors to the A institutional transactors.	
W is an AxA matrix of income transfers between the A institutional transactors.	Data can be taken from the IO table on import purchases from the External sector by each of the institutional transactors (except 'Corporate', where import purchases are already recorded for each individual production sector in **V**) and net commodity taxes paid to Government (again, with the exception of 'Corporate'). Data requirements for the remaining elements of sub-matrix ***W*** are additional to the SAM.
X is an AxB matrix of factor income payments to each of the aggregate transactors based on factor services supplied.	The total factor income payments are the totals of the 'other value-added' and 'income from employment' rows of the IO table. Where households provide all labour services to production, payments from Labour (L) to Households (H) are the entry in the (L) column of sub-matrix ***X*** . 'Other value-added' (payments to capital) are allocated based on the ownership of firms by the three aggregate transactors, Households, Government and Corporate. Shares of 'other value-added' accruing to foreign owners are transferred from Corporate to External in sub-matrix ***W***.
Y is a BxI matrix of payments to value-added/factors of production by each of the I production sectors.	Sub-matrix Y is identical to the 'income from employment' and 'other value-added' rows of lower left quadrant of the IO table.

The 1998 Scottish 3 sector SAM, based on the schematic SAM used in the AMOS modelling project at the University of Strathclyde, is shown here in the Excel file SC_SAM98.xls. The SAM itself is found in the worksheet named "SCOT_SAM_98". We shall discuss this SAM further a little later. It is available here just in case you wish to compare an actual SAM with the schemes shown above. We have put the SAM model together with the IO table from which much of its data is derived. The IO table is to be found in the sheet named "3SECTOR". In the SAM, the highlighted cells show the parts of the SAM that require data additional to that found in the IO table.

In a separate Excel workbook named SAM98_BLOCKS.xls, we have colour coded areas within the SAM itself to show you how the elements T, U, V, W, X and Y map into various blocks of the SAM table. You might like to view this document.

2. Balancing Identities

The balancing requirement of the SAM is that for each individual production sector, i, **and** for each institutional/aggregate transactor, a, **and** for each factor of production, b, the following identity must hold:

Total income (receipts) = total expenditure (outlays)

3. SAM Data Collection And Organisation: Construction Of A Set Of Income- Expenditure Accounts

The additional data requirements for the SAM (additional to those in the IO table, that is) are mainly concentrated in sub-matrix ***W*** in the schema shown above: transfers of income between the 5 main "transactors":

- Corporate sector(C)
- Government sector (G)
- Households (H)
- Capital Formation (CF)
- External sector (E)

The only other additional piece of information required is the shares of Other Value Added going to the various **local producer groups** in the sub-matrix ***X*** of the schematic SAM:

- corporate (privately operated firms)
- government (publicly operated firms)
- households (firms operated solely by self-employed owners)

Since all the entries that require additional data to that provided by the IO table are contained within the rows and columns of the five aggregate/institutional transactors, it is possible to deal with all these data gaps by constructing a set of income-expenditure accounts for each transactor.

On the following page, we show a template for the income-expenditure accounts used in the Scottish SAM developed at the University of Strathclyde, together with the estimated accounts themselves.

4. Additional Data Requirements for CGE Modelling : Investment Demands And Labour Supply Data

The IO table and SAM contain information on which sectors' outputs are used for the purposes of capital formation. However, what they do not tell us is which sectors the demand for this capital formation comes from.

IO tables tend to also report only the **full time equivalent** (FTE) employment by sector.

- This is sufficient labour market data for a demand-driven general equilibrium model like IO.
- However, for a more flexible CGE framework with an active supply-side, more information is required on supply conditions in the labour market.
- We also require data on the structure of the aggregate labour market, such as base year working age population, participation rate and unemployment.

In a long-run equilibrium (which the base year database is assumed to represent):

- the total amount of labour demanded must equal the total amount of labour supplied;
- the total amount of labour supplied is equal to the total labour force minus the number of unemployed;
- the total labour force is equal to the working-age population minus non-participants.

5. SAM databases for CGE modelling.

As noted above, the SAM framework is similar to a set of an IO accounts: as in the IO table, row entries in a SAM record incomes/receipts and column entries record expenditures/outlays. They are both examples of single entry book keeping, with every entry appearing in both a row and column (i.e. each item of income to one transactor must be an item of expenditure by another so that total receipts equal total outlays). However, unlike an IO table, a SAM does not just record income from, and expenditure on, the sale and purchase of goods and services. A SAM also records transfers of income (such as flows of profits to and from the operators to the owners of capital and other property, income taxes, social security payments etc.) between the five main aggregate transactors/institutions:

1. Households **(H)**
2. Corporate (all private firms) **(C)**
3. Government (as a producer and final demander) **(G)**
4. Capital Formation **(CF)**
5. External (UK , REU and ROW and tourists) **(E)**

Extension of Multi-Sectoral Modelling: Computable General Equilibrium Analysis and Modelling

In this lesson, we continue our discussion of general equilibrium modelling by looking at the following matters:

- CGE model calibration and specification
- An example of a CGE model: the single region, 3-sector AMOS model for Scotland, 1998
- Applications of CGE modelling
- Conclusions

Introduction to CGE

1. General remarks

Computable general equilibrium (CGE) models are also known as applied general equilibrium models (AGE) models. We shall use the former name in these notes. The acronym CGE implies that such models are:

- **Computable**: these models cannot in general be solved algebraically but they can be solved "numerically" using computational algorithms - these are sophisticated trial and error methods that search for solutions by trying initial sets of possible solutions values, and adjusting these values until a solution is obtained.
- They are **General equilibrium** models.
    - That is, a solution to such a model is one in which there is simultaneous equilibrium for **all** transactors and for all markets.

'General equilibrium' is in contrast to 'partial equilibrium', where equilibrium is studied in one market in isolation from all others. A difficulty with partial equilibrium analysis, though, is that changes in one market will often have knock-on effects on others markets. These knock-on effects cannot be studied in a partial equilibrium context. General equilibrium - by virtue of studying all markets jointly - avoids this potentially serious shortcoming.

- They are **models** of the economy

What is an economic model?

To get some insight, read (at least the first few pages) of this paper by Hal Varian of Berkeley.

The Wikipedia entry on this topic is also very good.

2. How do CGE models relate to standard IO models?

Recall the following characteristics of IO models:

- IO implies "only demand matters" (or only supply matters) - but not both
- IO only deals with quantities or prices
- IO assumes universal Leontief technology
    - No possibility of substitution in production or consumption *even if prices change* (in the price version of IO)

In contrast, CGE models relax these assumptions and typically allow:

- Non-passive supply-side: "supply matters" - usually **as well as**, not instead of, demand
- Relative prices become endogenous
- Transactors are allowed to respond to relative price changes - through **substitution** in production and demand (so we need to get rid of Leontief technology across the board)
- More flexibility, e.g. in technology and so non-linearities
- **More economic theory-consistent modelling**

3. Generic aspects of CGE models

If conditions are such that supply genuinely is passive (e.g. massive excess capacity, such as Keynes wrote about or during the Great Depression in the interwar USA) then

- CGE will generate IO results (aggregative structure - i.e. no price changes, increases in all quantities)
- IO models are in fact a special case of CGEs, with simple supply and technology assumptions and so linearity

CGE modelling is more difficult to characterise than IO models:

- CGE models are characterized by a huge variety of model types
- There is heterogeneity of views of the supply-side, and of technology, of the economy

Nonetheless, it is possible to identify some common features:

- Theoretical "roots" in Walrasian general equilibrium theory
- Elements of model construction
- Motivated by an attempt to shed light on key aspects of **economic policy** - hence inclusion in the MSc EMP certificate stage core classes.

4. Common theoretical antecedents

- The schematic of a typical CGE is reminiscent of the simplest macro-economics "circular flow of income" diagram. But within this general framework, CGE models are often characterized by
    - disaggregation of households and industries/commodities
    - all commodity and factor markets clear simultaneously - there is no excess demand or supply in any market in a general equilibrium
    - factor prices and employment levels are just sufficient to generate incomes and demands that equal commodity supplies

- **In the early developments of formal CGE theory**, modellers initially:
    - assumed universal perfect competition
    - focused on providing a proof of the **existence** of equilibrium in exchange and then production economies
    - but existence of equilibrium *per se* is really only the first step

- Scarf's 'constructive' proof and associated algorithm
    - Proof of existence also suggested as a means of solving a numerical/empirical version of a GE model
    - Can then solve for equilibrium - and changes in it brought about by, e.g. policy changes
    - Feasibility of numerical versions of GE models

- **Now, though:**

  - The existence of equilibrium is typically taken as given (we simply assume that there is a general equilibrium solution to our model, and proceed straight on to find that equilibrium)
  - CGE models are solved routinely, without Scarf-type algorithms
  - CGE models are not restricted to conditions of universal perfect competition: current applications typically use frameworks that encompass various degrees of imperfect competition
  - CGE models are used for modelling applications in very diverse areas, including economic development; monopolistic competition; and public economics.

- CGE requires that one can find a simultaneous equilibrium in all markets
  - CGE models focus on flows rather than stocks. Often initial stocks (of such things as physical capital and labour force) are not even identified in CGE models, as concern is placed on changes in stocks (i.e. flows) rather than the levels of stocks as such.
  - Hence, an important concern of CGE models is the source of supply and demand in each type of market of interest in the model
- CGE models can be complex
  - many goods and factor markets;
  - other transactors: government, rest of the world (ROW);
  - possible imperfections in markets.

CGE model specification

Any CGE model can be characterised in terms of a number of broad stages. These comprise (but not necessarily in the following order):

Stage 1: Dimensionality and specification

As with IO models, we must decide what sectors/activities/transactors we want to identify in the model.

Stage 2: Closure rule

IO models are "closed" simply by requiring that a solution must satisfy a set of accounting or balancing identities. These identities are brought about by supply adjusting passively to demand.

In CGE models, we must **make a choice** about the form of macroeconomic 'closure' that will be used. These closures sometimes deal with **how wages are set in labour markets** (for example, are they set at the regional or national levels?). Closure rules may also relate to **government fiscal rules or constraints** (for example, do we assume that the government budget must always be balanced, or do we assume that there is no constraint forcing the government budget to be in balance?). Closure rules may also be about **balance of payment constraints** (if any) that we assume must be satisfied. And there are many other behavioural assumptions or constraints that we might wish to examine that will generate other forms of macroeconomic closure.

- Development models often assume 'Keynesian closure' and involuntary unemployment (large excess capacity)
- Many models *assume* full employment - e.g. NAFTA studies, and other multi-country CGE models

- Closure matters, because **what you assume about the closure of the CGE model determines the nature of the results it will generate**
- Models may have range of alternatives, allowing the use to choose whichever he or she thinks is appropriate, or allowing the user to compare results under alternative macro-economic closure assumptions.

Which closures will we examine?

In the examples of CGE modelling discussed below, and in Briefing Note 4, we will consider only two of the many possible macroeconomics closures:

1. A bargaining real wage model
2. A national bargaining model

In these two cases the closure depends on how we believe the wage rate is being set in labour markers. Our modelling examples will also focus on the case of Scotland (the 'Regional Economy') within the context of the United Kingdom (the 'National economy'). But it could equally well be taken as one example of a more general template in which our economy of interest is a 'regional' economy within some broader 'aggregate' context (such as an EU member state within the EU, or a US state within the USA, or one national economy within a 'regional trading block').

Given this, in the Bargaining Real Wage closure, we assume that the wage in the regional economy of interest (Scotland, in our example) is set by a bargaining process between suppliers and demanders of labour in that regional economy.

In the second closure we example (the National Bargaining model), wages are assumed to be set by bargaining at the national (wider economy) level (the UK in our examples), and so are exogenously determined from Scotland's point of view.

More details on these closures will be given as we go along.

Stage 3: Assumption made about degree of competition

- Perfect competition still common in public economics
- Monopolistic competition and returns to scale (Dixit/Stiglitz) commonly assumed in trade models

Stage 4: If the CGE model we are using is to be dynamic, what form do the dynamics take?

- CGE models can take account of the passage of time, and tell us something about how long adjustments take to become fully completed. Recall that IO models were entirely silent about this.
- There are two ways of doing this:
    - **Intertemporal Optimisation**: this involves deducing how agents will behave optimally over time to maximise their lifetime utility, given inter-temporal budget constraints, and the specifying the equations within the CGE model so that they are consistent with this intertemporally optimal behaviour. Intertemporal optimisation-based CGE is relatively complex but many CGE models do adopt this route
    - **Recursive Dynamics**: stocks are updated from period to period given the flows that occur in each period within the model. For example, if wages rise in the local

economy, there might be inward migration of labour from outside the region. Our labour supply equations will specify a rule which states how much inward migration will take place for any given wage differential. So, in any given period, a wage differential will induce a particular amount of migration. The labour supply stock is then updated to take account of this flow, and the model then moves on to its calculations for the next period. This is a far simpler approach, and can give satisfactory results but the results do not necessarily reflect fully rational behaviour. The AMOS model used in our examples (and Briefing Note 4) is a recursive dynamic model

Stage 5: Spatial specification

- Single country/region (as with the AMOS model you are looking at)
- Multi-country/region (that is, inter-regional models), as in many trade models.

1. Database construction & collection of base year data

For CGE models, the database typically consists of a social accounting matrix (SAM) for the economy we seek to model.

As we showed earlier, the SAM builds on the data contained within an IO table, but more data are required, such as on the supply-side of the labour market (total population, participation rate; working age population), and on investment and capital stocks (investment demand; capital stocks; capital depreciation).

We also need to construct base year prices:

- The prices of labour, capital
- For some prices, we index them to unity (1) for convenience. This is justified because in a CGE exercise we are concerned with relative prices rather than the absolute levels of prices.

If one is building an intertemporal CGE model, time series data are required on some of the variables of interest. But this is not required for a recursive dynamic SAM, such as the AMOS model of Scotland.

2. Data are also required to inform parameters (calibration)

With IO modelling, all that is required is the A-matrix (and employment levels if one wishes to do employment extensions of IO). But CGE modelling is more rooted in economic theory and so this leads to more demanding data requirements.

For example, it is conventional to assume that households are utility maximizing and that firms are profit maximising/cost minimising and to specify and calibrate the model accordingly.

- We need to specify (either on the basis of *a priori* theory or by means of statistical estimation) the functional forms to be used in the various behavioural equations of the model, and the parameters of those relationships (recognising that functional forms may be non-linear in a variety of ways)
- There is a common use of hierarchical production and consumption structures (along with associated separability assumptions).

Hierarchical (or multi-level) production structures

Unlike input-output models, which are entirely based on Leontief technology, neoclassical theory guides specification of production in CGE models. In consequence, the CGE model does not represent factor demands as linear functions of output. Instead, factor demands depend on both output and relative prices. There is an exception, however, in relation to treatment of those goods and services that are used as intermediate inputs. In some circumstances, one might wish to use Leontief input-output production function where it is appropriate to represent production of output with fixed proportions of primary composite factors and composite intermediate inputs.

The composite primary factors generally enter the production process in a manner allowing factor substitution. Thus, production is best described as a hierarchical, multi-level or nested production process. Note that all factors in a constant elasticity of substitution (CES) function have the same elasticity of substitution between any pair of factors. To allow for differing elasticities between sets of factors, multi-level or "nested" production function forms are used in CGE, with each level containing a different set of factors and their own corresponding elasticities of substitution. That is, the use of a multi-level structure allows for use of both fixed-coefficients and price responsiveness in the CES form.

Source: Adapted from Computable General Equilibrium Modeling for Regional Analysis (Eliécer Vargas, Dean Schreiner, Gelson Tembo, and David Marcouiller), available via the internet at http://www.rri.wvu.edu/WebBook/Schreiner/chapter3.htm

3. Three ways parameters are determined

1. **By base year data** (e.g. labour intensity = sectoral FTE employment /sectoral output) ▫ structural parameters (change with structure of economy)
2. **Exogenously imposed -** here we use values obtained prior to the construction of the SAM. These will either consist of parameter values that have been econometrically estimated using time series or cross-section data (or both together: panel data) or they will come from other information that comes from prior research by your CGE team or by other empirical researchers. Exogenously imposed parameters often include those related to elasticities of substitution in production, and price elasticities of demand.
3. **Calibrated -** run model with no changes to solve for all remaining unknown parameters. This will include, for example, intercept values of all functions.

Calibration is normally to the base year SAM

- Values of 'key parameters' are identified first, as described above, either by econometric estimation of individual relationships by the modeller or by using externally generated (secondary) estimates.

- Then all remaining parameters are determined through reconciliation to the base year SAM.

Ideally, one would econometrically estimate all parameters used in the the CGE model. But this is very unlikely to be done in practice, as there are usually insufficient data to do parameter estimation of every parameter.

The approach taken with the AMOS project was a mixture of econometric estimation, using prior information from external sources, and calibration to base year data for remaining unknown parameters.

4. Features of calibrated models : source of criticism

In doing calibration of parameters, one assumes no errors or omitted variables, and we in effect "estimate" parameters using just a single observation, that pertaining to the base year relationships. As a consequence, we are bringing in to our CGE model estimates of parameters that have been estimated at best very inefficiently, and possibly ones that are heavily biased. Moreover, by not using standard statistical techniques to estimate parameters with multiple observation data sets, one has none of the conventional statistical measures that help us assess the quality of our estimates, and so we are working blindly in this regard.

As said above, it is generally agreed that one should only obtain parameter values by calibration when it is not possible to obtain them using econometric estimation or by using prior information. Unfortunately, given current levels of data availability, it is not yet feasible in most circumstances to obtain complete econometric estimation of CGE models, and few 'pure' econometric models exist.

We can improve upon calibration where new information becomes available, perhaps replacing calibrated parameters by ones that have been recently estimated econometrically, or by new external research findings. (One would then, of course, have to recalibrate the model for the remaining unknown (but now smaller) set of parameters.)

Given the limitations imposed by the need to calibrate parameters on single observations, it is important that one's use of a CGE model should be consistent with those limitations. Thus, some people use CGE models for forecasting purposes. This seems to be an inappropriate use, particularly if one is seeking quantitative rather than qualitative forecasts. It is better to work on the premise that model results are indicative of the qualitative nature of responses to policy changes, shocks and the like. The results tell us about what kinds of things will change, the likely directions of change, and their approximate magnitudes. The results should not be taken as reliable and precise point estimates of changes in the variables of interest.

5. Model solution

Solving CGE models involves the use of complex algorithms: these can either be constructed in-house or purchased off the shelf, such as the software called GAMS. Computationally it is much more difficult than what is required to solve IO models. As you know, even simple spreadsheet packages such as Excel can be used for IO models, because the linearity of IO relationships means that all that is required is to obtain the $\mathbf{[I-A]^{-1}}$ Leontief inverse matrix.

Analysis using the CGE model

A common motivating feature of CGE models is analyzing and evaluating the impacts of policies. But CGE analysis also considers other, non-policy, disturbances, such as the energy CGE models constructed partly in response to the energy crises in the 1970s.

Policies

Policies that are particularly amenable to CGE modelling include

- regional, environmental, development and structural policies

Effects

CGE analysis can be used to study the effects of any policy change or disturbance on such things as

- the economy itself, energy use, the natural environment
- the overall welfare impacts of the policy or disturbance

Model dimensionality and specification

Model dimensionality and specification will depend both on the target economy and the ***type of problem(s) it is designed to analyse***

e.g. AMOS programme (see 'AMOS' document). Now incorporates an emphasis on energy.

Application: the single region, 3-sector AMOS model for Scotland, 1998

Reading - for fuller details about the points made on this page see: AMOS (A Macro-Micro Model of Scotland)

1. Dimensionality/model specification

The initial 3-Sector AMOS framework - and the one we use here for teaching purposes - incorporates:

3 transactor groups	• households • firms • government
3 commodities and	• manufacturing (M) • non-manufacturing traded (NMT)

activities	• non-traded/sheltered (NT)
2 exogenous external transactors	• the rest of the UK (RUK) • the rest of the world (ROW)

There are four main components of final demand:

(Household) consumption	Treated as a linear homogenous function of real disposable income.

Investment	Treated in various ways, dependent upon the time frame that is chosen:
	
	• In the short- and medium-run the capital stock and its sectoral composition are fixed so that, even where investment is endogenous, capital stocks are not updated.
	    ○ in the short-run, both labour market and capital market are constrained so that stocks do not change
	    ○ in the medium run, the labour market starts to relax (so that labour supply can change) but the capital market has not yet begun to do so (and capital stocks remain fixed)
	• In the long-run, both labour and capital market constraints are fully relaxed. Equilibrium investment for each sector is endogenous and equal to depreciation with sectoral capital stocks set at their desired, cost-minimising levels.
	    ○ In the multi-period variant of the model, we let the model run forward over many periods. Investment in each period is equal to depreciation plus some fraction of the gap between actual and desired capital stock.
	• Note that short-run, medium-run and long-run are **conceptual** time frames, defined in terms of which constraints, if any, are in operation. They do not correspond to particular spans of real, calendar time.

Government expenditure	• The model can specify government expenditure in either of two ways:
	    ○ It is taken to be exogenous, or
	    ○ It can be treated as endogenous, linked to changes in income tax.
	
	Which is chosen will depend on the form of macroeconomic closure we assume in our simulation or policy analysis experiment.

Exports	Exports and imports are determined via an Armington link. That is, we specify a positive price elasticity of demand for

	exports, making them relative price sensitive. The Armington trade substitution elasticity is determined by the model user (or informed by econometric work). The way that exports are specified in the model means that we can examine the impacts of a pure export final demand shock, and in doing so will observe a gradual adjustment process. As export demand rises because of the shock, in the short and medium run prices will rise thus choking off some of the rise in export volumes. There is a partial initial adjustment of exports to the demand shock, and exports converge gradually to their full long-run equilibrium level only when (and if) there are no longer any price rises that serve to choke off export demand.

The model user can choose between perfect and imperfect competition and between different macroeconomic and labour market closures.

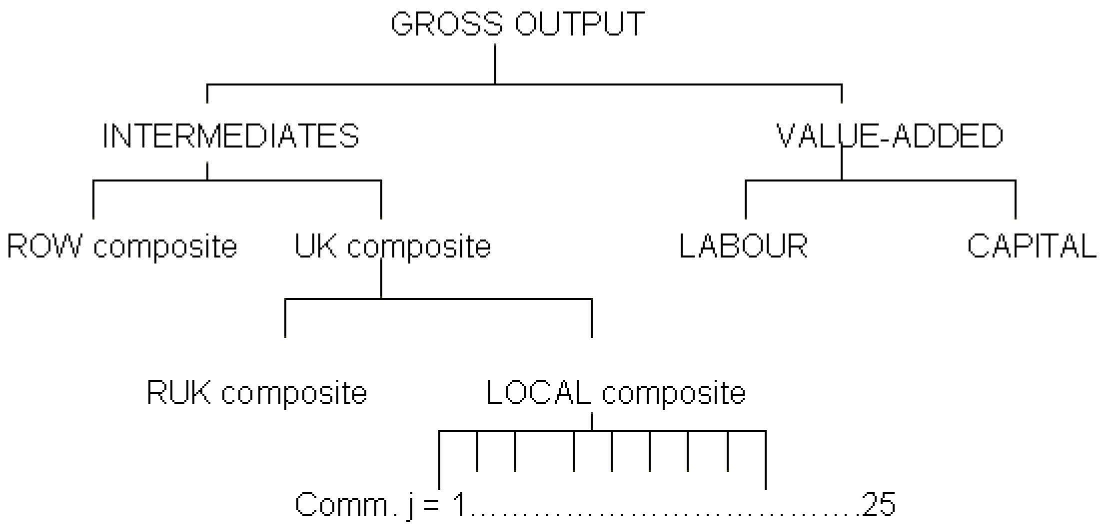
2. Database

1998 Scottish IO tables (aggregated to 3 sectors), augmented with other data for the SAM.

3. Parameterisation/calibration

Production and consumption structures fixed to one option to date, but the option varies across application. As stated earlier, production is determined by a hierarchical or multi-level production function.

Figure 1 : Production Structure In The Basic (3-sector) AMOS Framework



- The model specifies multi-level production functions
- Generally, these are of constant elasticity of substitution (CES) form, so there can be input substitution in response to relative price changes
- But we have Leontief and Cobb-Douglas (CD) technologies available as special cases. Leontief has elasticity of substitution equal to zero, whereas Cobb-Douglas has elasticity of substitution equal to one.

- In the CES functions, elasticities of substitution, σ, as with all parameter values, can be set for individual applications according to econometric or 'best guess' estimates.

- The production inputs are labour (L), capital (K) and intermediates (J), with a choice between locally produced intermediate commodities and imports from RUK or ROW.

This hierarchical structure allows for substitution between inputs.

For example, look at the block in which Value Added depends on two inputs, labour (L) and capital (K). If this part of the production structure were specified to be Leontief (so that Value Added can only be produced using fixed proportions of L and K, then even if relative prices of K and L changed, there could be no substitution of one input for the other.

But in a CGE model, Value Added will be produced not by Leontief technlogy but through a neo-classical production function in which the chosen input combination will depend on relative prices of K and L. Thus if K were to become more expensive for any reason, some K would be substituted by additional L (with the amount of substitution taking place depending on the sizes of the relative price change and the parameter governing the elasticity of substitution between the two inputs).

Cost minimisation

In all model configurations cost-minimisation in production is imposed, and the model solves for input choices that minimise production costs, subject to whatever constraints may be in operation.

4. Prices of intermediate goods

The prices of the intermediate goods that make up the intermediate composite are required to make our production structure work.

- All local input prices are endogenous to the system
- All import prices are exogenous (because they are externally-determined prices, and are not affected significantly by changes in demand from a small economy such as Scotland. Clearly, this assumption would not be valid for a 'large' open economy such as the USA).

The precise nature of the intermediate composite can change. Its composition will depend on relative prices of individual elements of intermediate input, and the possibilities for substitution between different sources and types of intermediate input at each level.

5. Wages

The precise form of the wage equation depends on what type of labour market regime is assumed to exist. One of the key features of the AMOS framework is that it incorporates five alternative labour market closures. This feature was chosen because the model-builders wished to examine policy changes under various alternative wage-setting regimes. Therefore, while this feature is central to AMOS, it need not be prominent in other CGE models. The general point here is that the options chosen for macroeconomic closure of any CGE model will depend on what particular questions that model was built to address.

The specification of the wage equation in each labour market "closure" is fairly standard:

1. Neo-classical/continuous market clearing	Here the wage adjusts so as to equate labour demand and labour supply
2. Keynesian	The nominal wage is exogenously determined at the regional level. The motivation for this would generally be a **national** bargaining regime. The aggregate labour supply function is suspended up to full employment.
3. Real wage resistance	The real wage is fixed - i.e. the nominal wage is a mark up on the consumer price index.
4. Exogenous labour supply	A fixed proportional relationship exists between employment and working population (this is often

	taken to be the closure for national CGE models).
5. Regional wage bargaining (also referred to as the bargained real wage, BRW, closure).	The regional consumption wage is directly (positively) related to workers' bargaining power and inversely (negatively) related to the regional unemployment rate via a bargained real wage function. Note that this closure does imply that local wages are flexible in that they respond to the local excess demand for labour.

The two that we shall focus on in this class have been highlighted in red in the table above.

6. Dynamics

AMOS is not an intertemporal optimisation model, but it can be run as a recursive dynamic model. (in the exercises we set for your Briefing Notes, we shall use results for a sequence of successive periods generated by this recursive dynamic method).

Labour and capital stocks are updated period by period. Capital stock updating takes place by means each sector's capital stock being updated between periods via a simple capital stock adjustment procedure. This specifies that investment equals depreciation plus some fraction of the gap between the desired level of the capital stock and its actual level.

7. Labour market/population

The regional economy is initially assumed to have zero net migration (as in the short-run in the static model), and ultimately, net migration flows re-establish a long-run population equilibrium.

- Net in-migration in any one period is taken to be positively related to the regional real wage differential in the target region relative to the national economy and negatively related to the regional unemployment rate.
- This migration model is based on that in Harris and Todaro (1970).

More material to be added in here in a later update.